A COMPARISON OF THREE TEACHER EVALUATION METHODS
AND THE IMPACT ON COLLEGE READINESS

Tamy L. Smalskas, B.A. B.S. M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2013

APPROVED:

Jimmy Byrd, Major Professor
Mark A Davis, Minor Professor
John Brooks, Committee Member
Linda Stromberg, Committee Member
Nancy Nelson, Chair of the Department of Teacher Education and Administration
Jerry D. Thomas, Dean of the College of Education
Mark Wardell, Dean of the Toulouse Graduate School

Smalskas, Tamy L. A comparison of three teacher evaluation methods and the impact on college readiness. Doctor of Education (Education Administration), December 2013, 135 pp., 37 tables, 1 illustration, references, 110 titles.

Much attention in recent years has gone to the evaluation of teacher effectiveness, and some scholars have developed conceptual models to evaluate the effectiveness. The purpose of this study was to compare three teacher evaluation models – the Texas Professional Development Appraisal System (PDAS), the teacher index model (TI), and the value-added model (VAM) – to determine teacher effectiveness using student demographic and longitudinal academic data. Predictive data from students included economic disadvantage status, ethnicity, gender, participation in special education, limited English proficiency, and performance on Texas Assessment of Knowledge and Skills (TAKS). Data serving as dependent variables were scores from Scholastic Aptitude Test (SAT®) verbal/critical reasoning and mathematics. These data came from 1,714 students who were 9.7% Hispanic, 9.2% African American, and 81.2% White. The models were tested for 64 English language arts teachers and 109 mathematics teachers, using student examination scores from the SAT® verbal/critical reasoning and mathematics. The data were aligned for specific faculty members and the students whom they taught during the year of the study.

The results of the study indicated that the TI and VAM explained approximately 42% of the variance in college entrance exam scores from the SAT® verbal/critical reasoning and mathematics (R^2 = 0.418) across mathematics and English language arts teachers, whereas the TI model explained approximately 40% of the variance in the

SAT® scores ($R^2 = 0.402$). The difference, however, in the *R*-squared values between the VAM and the TI model was not statistically significant ($t\ (169) = 1.84$, $p > 0.05$), suggesting that both models provided similar results. The least effective model used to predict student success on college entrance exams was the PDAS, which is a state-adopted model currently in use in over 1,000 school districts in Texas, The teacher PDAS scores explained approximately 36% of the variance in student success on the SAT® ($R^2 = 0.359$).

The study provides school leadership with information about alternative methods of evaluating teacher effectiveness without difficult formulas or high costs associated with hiring statisticians. In addition, results indicate that the models vary significantly in the extent to which they can predict which teachers are most effective in preparing students for college. This study also emphasizes the critical need to provide teacher evaluations that align with student achievement on college entrance exams.

ACKNOWLEDGMENTS

Completing this dissertation and graduating with my PhD would have not been possible without several amazing and supportive individuals in my life. I want to begin first by thanking God for giving me the strength and ability to make it through the past few years and especially during a very trying time with my family for two of the years. I need to thank my two baby girls; Dakota and Darbi, for allowing mommy to take class and work on homework. I am done and am 100% yours. I also want to thank my husband for his patience and understanding during times of frustration, plus all the days, nights and weekends he watched the girls so I could complete assignments or work on my dissertation. My mom and dad, Allan and Wendy Midtdal provided me with encouragement and the values and integrity of being a doctoral student and thank them for their love and encouragement. I need to sincerely thank my school district, Mary Clark, Dr. JD Kennedy, Geoff Sanderson, and Jason Bird for allowing me to complete the study and utilize the district data to accomplish my goals. I need to acknowledge and thank my editor; Patti Dale for great feedback and a motivator. My committee members; Dr. John Brooks, Dr. Linda Stromberg and Dr. Mark Davis all provided me excellent feedback and encouragement throughout the process. To my dearest friends that have walked through this journey with me, encouraging me along the way; Chaurcley, Deb, Abby, Marion, Karen, Melissa, Cindy and Hallie. I love each of you for your constant encouragement! Finally, I could not have done this without the support and knowledge of my advisor Dr. Jimmy Byrd. Dr. Byrd provided constant guidance and direction for this entire process by keeping me on target and especially providing the support for all the data analysis in this study. Huge thank you to Dr. Byrd!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Measuring the impact of teacher effectiveness on student achievement is a continuing challenge for public education (Haushek & Rivkin, 2012). More concretely, teachers receive inflated appraisals for job performance, which often involve only one evaluation instrument, and the process may or may not be adequate in terms of determining how well students are being prepared for college (New Teacher Project, 2007; Toch, 2008; Chait, 2010; Schmoker 2012). In addition, there are few districts in Texas that go beyond the typical rubric checklist to evaluate teachers and policies are rarely in place to determine workforce productivity of individual teachers in terms of measuring student academic growth of one-year (Braun, 2005). Education is important to the global economy; for American students to effectively challenge a leveled playing field students must be intentionally prepared for post-secondary readiness by acquiring the requisite skills in high school (Fatima, 2009; Conley, 2010a; Hanushek & Rivkin, 2012).

Although many school districts utilize simple performance appraisals such as standardized testing, statistics indicate that schools are unsuccessfully preparing students with the necessary skills to complete college (Conley, 2003b). For example, fewer than one out of two students met the college readiness standards in math and verbal skills on the American College Testing (ACT®) and Scholastic Aptitude Tests (SAT®) in 2010, yet ineffective teachers are rarely dismissed or given a less than satisfactory evaluation (Chait, 2010; Smith, 2012). Teacher effectiveness is difficult to measure by simply using one evaluation method. Additionally, teacher evaluations do not align with strategies or criteria that provide all graduates the appropriate skills and knowledge to help them succeed beyond high school (Conley, 2010a, Chait, 2010).

While current research on teacher effectiveness indicates effective teachers impact the academic growth of students, students continue to graduate high school unprepared for college success (Sanders & Rivers, 1996; Darling-Hammond & Youngs, 2002; Hanushek, 2011; Conley, 2010a). Increasing the numbers of students with post-secondary degrees will improve the workforce, with an increase of approximately $70 trillion on the economic gross domestic product (GDP) (Fatima, 2009; Hanushek & Rivkin, 2012). Demonstrating a need for improving the measure of teacher effectiveness on preparing students for college, only 29.4% of students between the ages of 18-24 will graduate high school, and only 45.6% of that group will complete some college coursework or earn an associate's degree (U.S. Census Bureau, 2012). A method in which to ensure students are prepared for college is to equip high schools with effective teachers and provide multiple evaluation methods that accurately measure job performance (Chait, 2010; Hanushek & Rivkin, 2012).

College Readiness

State standardized tests are not an accurate indicator for teacher effectiveness based on college readiness standards (Conley, 2003a). Conley (2003a) examined 66 high school state assessments and determined that high schools are not effectively assessing students in a manner that correlates to success in entry-level college or university classes. Conley (2003a) infers that most states are not providing high school students with an accurate predictor of college readiness. A good measure of college readiness, are looking at student performance on the ACT® or SAT® (Hanover Research, 2013). The ability to predict college readiness, and having a tool to evaluate effective teachers, will provide administrators valuable information on job performance within campuses (Fatima, 2009).

Despite these inconsistencies, states continue using standardized test with a disregard for alignment of college-readiness skills. Conley (2010a) states “Given the tremendous variance in academic skills of high school graduates, it is no surprise that many struggle academically when they seek to advance their education beyond high school” (p. 6). Even taking high school classes that are described, as college-prep does not necessarily mean the student is prepared for college success according to Conley (2003b). Teachers are not provided with appropriate professional learning or guidance on the skills and knowledge students need to develop to ensure success for entry-level college courses (Conley, 2005). Conversely, students are graduating from high school with exaggerated information about their ability to begin and succeed in a college level course. Evaluations need to include an index to measure teacher impact on readiness for college as determined through workforce productivity skills (Thum, 2003; Conley, 2010a). A college education is the major determinant of employment earnings for American workers; public schools need to do a better job of evaluating teacher effectiveness on preparing students for college (Chait, 2010; Conley 2010a). Preparing students with college skills would increase students’ effectiveness upon graduating from high school and provide those students with the knowledge and proficiencies enabling them to be successful in obtaining post-secondary degrees (Becker, 1993, Fatima, 2009).

Teacher Evaluations and Student Achievement

Ineffective teachers can negatively affect the annual achievement growth of a student's learning by at least one year. This signals the importance of highly effective teachers for all students (Hanushek, 1992, 2003, 2011; Hanushek & Rivkin, 2010; Sanders & Rivers, 1996). However, connecting effective teaching and capturing it in an appropriate manner for teacher evaluations remains an issue, since studies of poor performing schools with ineffective teachers reveal teachers receiving high evaluation scores (New Teacher Project, 2007, Toch, 2008; Chait, 2010; Danielson, 2010; DeNisi & Sonesh, 2011). Furthermore, accountability measures beckon for the appropriate use of multiple forms of data to evaluate teacher effectiveness regarding college readiness and to align the job performance of teachers to the appraisal process (Fletcher, 2001; Cogshall, 2012). This premise suggests if teachers are doing a thorough job of teaching curriculum and preparing students for college, then evaluations should reflect job performance, and job performance will subsequently assess how effective teachers prepare students for college (Conley, 2010b; Marzano & Waters, 2009; DeNisi & Sonesh, 2011). Teachers' effectiveness cannot be captured solely by observations alone; it would be beneficial for campus principals to have the skill of identifying academic growth of students using some form of college readiness standard (Fatima, 2009; Hanushek & Rivkin, 2012).

In addition to poor evaluation systems of measuring teacher effectiveness, teachers need curriculum, resources, and professional development to incorporate the best instructional strategies in the classroom based on college readiness standards (Conley, 2010a; Marzano, Frontier, & Livingston, 2011; Coggshall, 2012). Student achievement correlates to the quality of the teacher (Hanushek & Rivkin, 2012), and teachers need appropriate resources and skills to provide the necessary skills and knowledge for students to be successful (Darling-Hammond, 2006; Conley 2005, 2010a). The success of student learning and overall development are dependent on the effectiveness of a teacher and instructional strategies used, resources implemented, and application of professional development (Marzano et al., 2011). Poor quality teaching or teachers who are not effective in the classroom can cost a student one year of academic growth (Hanushek, 1992; Sanders & Rivers, 1996). Hence, matching teacher performance with college readiness standards, in terms of increasing academic growth of students, should be captured in performance appraisals, providing teachers with feedback in areas to improve (Tziner, Murphy, & Cleveland, 2005). Performance ratings should be clear, relevant, inclusive of a measurement scale that is not questionable, and align with professional learning in order to address deficiencies of teaching college readiness standards (Kline & Sulsky, 2009; Conley, 2010a).

College Entrance Exams

There is a discrepancy between student achievement and performance ratings of teachers of low performing schools. Teacher effectiveness toward meeting college preparedness standards can be determined by using student results from standardized testing, such as the ACT® or SAT® (Conley, 2010b). The United States had only 1.6 million students take the ACT® test with only 25% of those students meeting all four ACT® benchmark scores (ACT®, 2012b). Only 67% of the same cohort was prepared to take college level classes. The results on the ACT® show a greater deficit in Texas with only 24% of the students meeting all benchmarks and only 61% meeting the English benchmark, far below the national average (ACT®, 2012c). In addition, a recent New York College Board news release published the importance of increasing rigor in American schools. The report indicated only 43% of the 2012 graduating class who took the SAT® "achieved the level of academic preparedness associated with a high likelihood of college success" (SAT®, 2012b, p.1) Therefore, capturing the data of college entrance exams such as the ACT® or SAT® can serve as another measure to validate teacher effectiveness while also being used as an indicator within the job performance evaluation (Conley, 2003b; SAT®, 2012a).

No Child Left Behind and Race to the Top

The reauthorization of PL-142, commonly known as No Child Left Behind (U.S. Department of Government, 2004), sought to improve teacher quality by requiring schools to hire teachers in core subject areas who are highly qualified (U.S. Department of Education, 2004. The term "highly qualified" is defined as having at least a bachelor's degree, a certification of the subject area, and demonstrated subject matter competence (U.S. Department of Education, 2012). NCLB (2001) further affirms that students are to have effective teachers with valuable skills and training, who can ensure their academic readiness for college. According to Hershberg (2005), NCLB moved in the right direction of having accountability on student achievement outcomes; however, NCLB fell short by holding schools, rather than teachers, accountable for learning outcomes.

The push towards focusing on measuring teacher effectiveness in terms of student achievement arrived when the government signed the American Recovery and Reinvestment Act (AARA) of 2009 (U.S. Department of Education, 2009). Developing a way to capture teacher effectiveness from using data from student achievement is the responsibility of school districts; subsequently, school districts are rewarded financially when they prove teacher effectiveness increases student achievement. The ability to measure teacher effectiveness is funded through the Race to The Top (RTTT) program. RTTT pedagogy gives school districts permission to evaluate a teacher's effectiveness as measured by student performance on college entrance exams as another method to improve job performance (U.S. Department of Education, 2009). According to the U.S. Department of Education (2009), RTTT defines an effective teacher in the following way:

> A teacher whose students achieve acceptable rates (e.g., at least one grade level in an academic year) of student growth**…**States, LEAs, or schools must include multiple measures, provided that teacher effectiveness is evaluated, in significant part, by student growth**…**Supplemental measures may include, for example, multiple observation-based assessments of teacher performance (p. 12).

RTTT shows a national effort to improve teacher evaluations, including ways in which to capture student academic growth in addition to providing the ability to determine teacher effectiveness, as measured by college preparedness of high school students (Conley 2010a).

According to the federal government's definition of teacher effectiveness, school districts must determine multiple ways to measure the academic growth of students. One such method to measure this growth has been defined as the value-added model (VAM), which uses a three-level hierarchical linear model (HLM) (Bryk & Raudenbush, 1988; Braun, 2005). The benefit of using VAM is its capability to isolate the impact of student learning based on instruction by the teacher. Likewise, VAM can provide information about annual academic growth of students attributed to the effectiveness of the teacher, classroom environment, school or district (Hershberg, 2005).

Problem Statement

Emergent research supports that high school students are not prepared for college after graduating (Conley, 2003a; SAT®, 2012a). In particular, there are a number of students who do not meet the benchmark scores on the ACT® or SAT® in mathematics or English (SAT®, 2012b). While the literature indicates there are mixed results of capturing the effectiveness of teachers, through the use of student achievement tests alone, it is neither known nor has any evidence been provided to measure teacher effectiveness in terms of preparing students for college (Hanushek & Rivkin, 2012; Conley, 2003b; Conley, 2010a; Darling-Hammond & Youngs, 2002).

Teacher evaluations need to include multiple methods of capturing effective teaching (Danielson, 2010; Becker, 1993, Hanushek, 2011). This includes providing methods of calculating effective teaching and using a college readiness indicator as the data source (Conley, 2010a). Equipping more students for college during high school, enables more students to graduate with skills and knowledge to earn post-secondary degrees (Fatima, 2009). The problem is that teacher effectiveness is not represented accurately through teacher evaluations, and teacher evaluations do not precisely reflect student achievement as determined by college readiness standards.

Theoretical Framework

The human capital theory is the foundation of this study (Becker, 1993). Human capital theory provides support for the argument that preparing students for college is important, as well as providing a framework for measuring teacher effectiveness and job performance. For the purpose of this study, human capital theory is defined as a form of capital investment in the labor force and economic growth through education and schooling (Becker, 1993). Mincer (1958) first explained the human capital theory and discovered “differences in training result in differences in levels of earnings among ’occupations’ as well as in differences in slopes of life-paths of earnings among occupations” (p.288). Becker (1994) points out that graduating from high school and completing a college education raises a person’s income, thus having a substantial impact on the economy of the U.S. Moreover, Nafukho, Hairston, & Brooks (2004) connect human capital theory to college and career readiness standards by establishing how investing in education and training is a development for individuals, communities, and the economy.

The research on college and career readiness by Conley (2010a) serves as the catalyst to support the framework of this study. State research on college readiness or is it college and career readiness?

Conley (2003a) indicates items on state achievement tests are not challenging enough nor indicate potential for college readiness. The author further suggests other forms of student data, such as the first year college grade point average (GPA), SAT® test scores, ACT® test scores, high school grades, and entry-level college courses can indicate potential for college success.

The four key components to college and career readiness include key cognitive strategies, key learning skills and techniques, key transition knowledge and skill and key content knowledge and are usually captured with college entrance exams (Conley, 2012). Through the analysis of student test scores, this study attempts to determine which forms of teacher evaluation are better predictors of college preparedness, as determined by ACT® reading and math and SAT® verbal/critical reasoning and math scores.

Teacher quality as reported by capturing effective teaching, is complex and consists of many variables. Previous research indicates teachers are the most important component for improving student achievement (Hanushek 1992, 2003, 2011; Rivers & Sanders, 1996; Darling-Hammond & Bradsford, 2005). This study elaborates upon prior research conducted by Hanushek (2011), who reported that the key element defining a school's impact on student achievement is teacher quality. Providing highly effective teachers will ensure students the opportunity to develop college readiness skills, and the measure of teaching quality will be represented by an accurate appraisal (Fletcher, 2001; Conley, 2010a).

Using the human capital theory as the underpinning as a foundation, along with a combination of educational research, this study seeks to establish conceptual framework of the impact of teacher effectiveness on college readiness. Measures were included by analyzing teacher and student variables, examining value-added growth models of achievement tests, measuring workforce productivity of teachers, and analyzing the Professional Development Appraisal System (PDAS) instrument to determine an effective measure of college readiness in terms of ACT® and SAT®

scores. The findings will contribute to an understanding of the contributions of public education to the human capital, which is essential for teachers, administrators and school districts to understand.

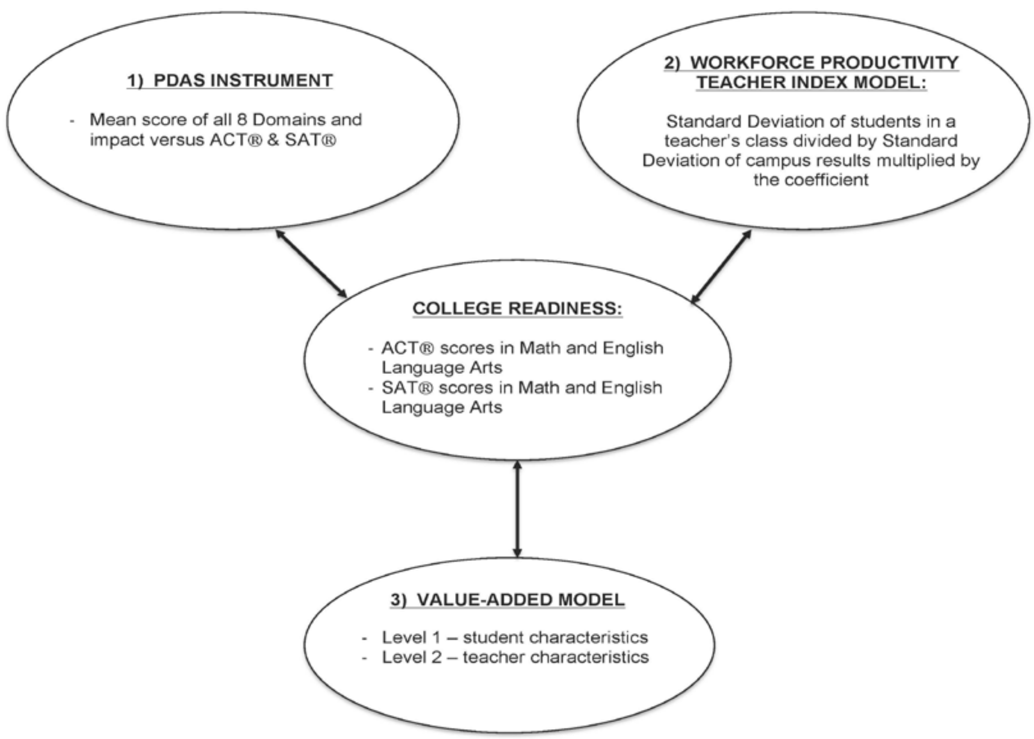

Figure 1. Conceptual framework: Teacher evaluations on college readiness.

Purpose of the Study

The purpose of this quantitative study is to compare three teacher evaluation methods to determine teacher effectiveness in terms of preparing students for college within a selected North Texas public school district. The dependent variables include scores from two college entrance exams, ACT® and SAT® in math and English. This research also includes comparing three different teacher evaluation methods as the independent variables. The independent variables analyzed in Model 1 include the evaluation scores from the PDAS. Model 2 is called the teacher index model. The independent variable in Model 2 is the standard deviation of scores on the Texas Assessment of Knowledge and Skills (TAKS) of students in a teacher's classroom divided by the overall standard deviation of campus TAKS scores. The third method uses a multilevel hierarchical model (HLM) to analyze teacher effectiveness controlling for multiple student characteristics such as gender, race, economic disadvantage status, limited English proficiency (LEP) and special education (Sped) status. In addition, teacher characteristics such as gender, race, degree earned, and total years of experience are controlled and analyzed. The three methods were analyzed to determine which evaluation tool is the better indicator of effective teachers in terms of preparing students for college success.

Research Questions

The following research questions guided this study:

1. Does the Professional Development and Appraisal System (PDAS) instrument predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?

2. Does the teacher index model predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?
3. Does the value-added model (VAM), used to derive a teacher effectiveness score, predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?

CHAPTER 2

REVIEW OF LITERATURE

This literature review provides current research on teacher effectiveness on student achievement, defines college readiness, provides a perspective on teacher evaluations, and investigates methods to analyze student achievement growth and the role the principal plays in determining effective teaching. This section also provides an in-depth look at the performance appraisal system from a human resource management (HRM) perspective. Additionally, the research of three evaluation tools; value-added model (VAM), professional development appraisal system (PDAS), and a teacher index model were examined. Together, these topics explore the framework of measuring teacher effectiveness and the impact on college readiness as measured by student achievement of high school students.

Teacher Impact on Student Achievement

Teachers are the primary contributor to students' academic achievement and development (Goe & Croft, 2009). Since college success can be predetermined by knowledge and skills learned in high school, evaluations should accurately reflect teacher effectiveness on college preparedness (Conley, 2010a). As public education faces the mandates from No Child Left Behind (NCLB) and all the state accountability programs, defining teacher effectiveness, characteristics of teacher quality, or teaching quality is a daunting task (Looney, 2011). NCLB expects all states and school districts to ensure highly qualified teachers. However, simply defining teachers as highly qualified does not necessarily mean they can effectively deliver quality instruction or have any positive influence on student achievement in terms of college readiness (Goe, 2007; Conley, 2008). There are problems in defining what constitutes teacher quality along with the ability to measure teacher effectiveness through the use of teacher evaluations (Chait, 2010).

Research shows (Schmoker, 2006; Hanushek & Rivkin, 2012) a significant causal relationship between teacher effectiveness and student achievement. Because of this, it is important to be able to retrieve information about teacher effectiveness and use that data to make decisions regarding teacher performance and outcomes.

An empirical research study of students in Grades 2 through 6, Hanushek (1992) discovered extremely substantial differences among the quality of teachers, indicating teachers have a definite effect on student achievement. The results of this study estimated the difference between having an excellent teacher and having a poor teacher as one grade-level discrepancy in test performance. A similar study conducted by Sanders and Rivers (1996) discovered that moving a low achieving student from an ineffective teacher to an effective teacher increases that student's chance of gaining back academic knowledge. Teachers who were in the bottom quintile of teacher effectiveness are ineffective with high-performing and low-performing students. Teachers in the fifth quintile (the highest effectiveness) are effective with all student ability levels. Therefore, effective high school teachers can increase college readiness skills of a student, regardless if the student is high performing or low performing (Sanders & Rivers, 1996).

Many high schools in Texas are also concerned about ensuring students graduate on time and with a particular cohort (TEA, 2005). To maintain graduation rates, principals and school leaders need to make sure students are not going to fall behind. Babu and Mendro (2003) content that when effective teachers taught low-achieving fourth-grade students for three years in a row, those students were twice as likely to pass a seventh-grade math test. Conversely, students who were low achievers assigned to ineffective teachers for the same amount of time did not see similar results. This longitudinal study conducted in a large urban school district indicates that having a teacher considered detrimental, in terms of academic growth, impacts both student achievement for the next two years as well as graduating on time (Babu & Mendro, 2003). Closing this achievement gap becomes the challenge for high school principals. When students are not successful on meeting minimum scores on state standardized tests, it is critical to assign effective teachers to poor performing students. According to Hanushek (2011), quality teachers and effective teaching has 6 to 10 times as much impact on student achievement than any other factors combined. This indicates the importance of measuring teacher effectiveness to ensure proper placement of struggling students with quality teachers.

Defining Teacher Effectiveness

NCLB created a framework of defined requirements for highly qualified teachers (U.S. Department of Education, 2004). One requirement is to ensure teachers are fully certified; in addition, each state must provide the criteria for being fully certified (U.S. Department of Education, 2009). However, teachers who hold certification based on such criteria do not necessarily have a positive impact on student learning (Hanushek, 2003; Hanushek & Rivkin, 2012). In addition to connecting teacher effectiveness to student achievement, it is important to define the measure of a “highly qualified” teacher. While there are various definitions and characteristics of what defines teacher quality or effective teaching, there is no clear consensus as to what contributes to teacher effectiveness (Goe, 2007). One focus area that determines teacher effectiveness is the impact teacher certification has on student achievement. Darling-Hammond & Youngs (2002) concur that teacher effectiveness is attached to student achievement; however, effectiveness of a teacher cannot be connected directly to teacher certification. Other studies conducted, however, support that teacher certification can have significant impact on student achievement, but teacher credentials do not predict quality teaching (Goldhaber & Anthony, 2007). Clotfelter, Ladd and Vigdor (2007) found that teacher credentials, as the measure by degree earned, have significant impact on math achievement, but no significant correlation on reading performance. In addition, a teacher who may have a national board certification, license, or a master's degree is not necessarily more effective than a teacher who has more experience and only a bachelor's degree (Clotfelter et al., 2007).

With this in mind, measuring teacher effectiveness, as determined by student success in college, cannot be determined by a lone definition of "highly qualified": a certification, degree obtained or a state license (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012). This discrepancy in research about teacher credentials can be confusing to school administrators, especially when searching for effective teachers (Looney, 2011).

Effective Teaching Characteristics

Research has sought to identify teaching attributes and instructional strategies that influence student achievement. Teachers who demonstrate higher results on student achievement tests possess stronger pedagogical knowledge than their less effective peers. For example, mathematics teachers with a strong pedagogical background provide a full year or more of learning for students than students who had teachers with a weaker pedagogical background (Baumert, Kunter, Blum, Brunner, Voss, Jordan, Klusmann, Krauss, Neubrand, Tsai, 2010). This supports teacher-training programs where teachers receive better pedagogical training (Rice, 2003). However, other researchers have indicated that knowledge of content subject matter, verbal ability, and higher literacy skills show the most significant increase in student achievement (Darling-Hammond & Youngs, 2002; Rice, 2003; Namaghi, 2010). Teachers who can communicate verbally or textually and those who have higher cognitive ACT® scores, also have a greater influence on the achievement level of students, indicating additional measures for identifying teacher effectiveness on student achievement in terms of college preparedness (Ferguson & Ladd 1996).

Effective teachers also need to have good content knowledge of the subject they teach, including strategies and methods to meet diverse learning needs (Darling-Hammond, 2006). Teachers who have well-structured classrooms, routines, procedures, and strong classroom management skills, also show results of high levels of student achievement (Hattie 2009). These components are identifiable with classroom observations and can serve as another quantifiable tool for identifying teachers who effectively impact student achievement (Chait, 2010).

Teacher effectiveness can also be measured by teaching practices. Teachers, who have positive influences on student learning and demonstrate success in student achievement, exhibit the following teaching practices:

- Understand subject matter deeply and flexibly
- Connect what is to be learned to student's prior knowledge and experience
- Use instructional strategies that help student draw connections, apply what they're learning, practice new skills, and monitor their own learning
- Assess student learning continuously and adapt to teacher to student needs
- Provide clear standards, constant feedback, and opportunities for revising work
- Develop and effectively manage a collaborative classroom for all students (Darling-Hammond & Bradsford, 2005, cited in Darling-Hammond et al., 2012, p. 13).

Furthermore, Darling-Hammond et al., (2012) proposed that if teachers at the high school level increased rigor and expectations in the classroom, students would develop skills for success in college (Conley, 2010a).

Marzano, Frontier, and Livingston (2011) suggest campus principals need to provide effective teaching strategies and methods to improve student achievement. This model consists of four domains: 1) classroom strategies and behaviors, 2) planning and preparing, 3) reflecting on teaching and, 4) collegiality and professionalism. Each domain Marzano et al. (2011) suggests for improving student achievement contains additional elements, which reflect best practice research strategies and behavior to improve student achievement. Domain 1, utilizing effective classroom strategies and behaviors, is the most valuable domain that has a direct impact on student achievement, as it supports well organized and designed lesson plans. Additionally, Schmoker (2011) highlights three elements that improve student learning in every subject area:

1. What we teach. The simple coherent curriculum;
2. How we teach. The sound structures and importance of good lessons employing the basic elements, and;
3. Authentic literacy. This is the spine that holds all subjects together. Students need to be able to read and write at high levels.

Providing adequate professional learning on quality instruction practices increases student achievement. Hairrell, Rupley, Edmonds, Larsen, Simmons, Willson, Byrns, and Vaughn (2011) discovered when teachers were provided with curriculum-based professional development with “specific instructional strategies such as corrective feedback, instructional pacing, and level of student engagement” (p.254) fourth grade social studies teachers had huge gains in student achievement. In addition to professional development, teachers who work collaboratively with peers, know how to assess students, and use student assessments to provide timely, specific feedback show a higher effectiveness on student learning (Looney, 2011).

Research supports that effective teaching can be measured in terms of student achievement, and clearly defined elements that support quality teaching should be measured in teacher evaluations (Chait, 2010). Providing teachers with the knowledge and skills to utilize the above-mentioned strategies, then training principals on how to identify and assess appropriate strategies aligned with college bound skills, should provide improved college readiness in students (Kaplan & Owings, 2001; Conley, 2012). This philosophy must be developed within primary grades and up to secondary high schools. As King and Watson (2010) state, above all, “accomplished teachers embrace and express the goal of college readiness and the expectation of high school graduation and college success for each and every student beginning in prekindergarten” (p. 178). The practical implications of research can lead to increased student performance, support in preparing our students for college, and ensuring effective teachers for all students (Marzano et al., 2011; Conley 2010a).

College Remediation and College Readiness

Our Nation’s economic future depends on developing a highly skilled and educated labor force capable of competing with other nations. We know that adults with a postsecondary degree earn about 40% more than those who have just a high school diploma (U.S. Census Bureau, 2012). According to Schneider and yon (2011) our “states spend more than $1.3 billion per year on students who drop out during their first year of college and the federal government spends an additional $300 million per year” (p. 2). President Obama established a target of having the highest percentage of adults with college degrees in the world by 2020 (Schneider & Yin, 2011). This reaction is a result of a report by the American Institute for Research, indicating that “students who started college in the fall of 2002 as full time students, failed to graduate six years later. This issue further exacerbates by costing the nation approximately 3.8 billion in income, $566 million in federal income taxes, and $164 million in state income taxes nationwide” (Schneider & Yin, 2011).

In addition to the aforementioned statistics, Conley (2010b) analyzed thousands of entry-level college courses and discovered from the teachers and data collected that the most consistent concern regarding potential success in college is that students are not proficient in a range of cognitive key strategies. High school students have an inability to translate knowledge into cognitively complex tasks and organize these elements sequentially and progressively. Conley (2008), states the key elements of success in college are “the development of cognitive and meta-cognitive capabilities of incoming students: analysis, interpretation, precision and accuracy, problem solving, and reasoning” (p. 3). Student success in college also requires the student to establish and maintain self-management behaviors such as time management, strategic study skills, persistence, and recognition of performance, self-control, and use of study groups. These skills are far and beyond just having cognitive skills and can be tied to observable measures to incorporate into teacher evaluations (Danielson, 2010; Donaldson, 2011).

High school graduates experience tremendous differences between high school and college. College students are expected to act like adults, take responsibility to ensure they complete assignments, study appropriately and be aware when coursework falters. Students need to be prepared, as college professors are not always going to provide “extra credit” or “extra work” to improve grades (Conley, 2011; Fatima, 2009). College instructors often emphasize different aspects of the material being taught. Specifically, college content requires students to analyze conflicting explanations of phenomena, make inferences, interpret results, support their arguments with facts or evidence, solve complex problems with no obvious answers, reach their own conclusions, conduct research, and think more critically about what they are learning (Conley, 2008; Conley 2010a; Conley 2010b; Conley, 2011). The literature also states the most significant indicator of college readiness is the need for students to be able to read and write effectively. Conley (2003b), found college students write several papers requiring them to be able to read multiple chapters, books, and research papers, and then demonstrate the ability to organize, reason, and write with evidence from research. For this reason, students leaving high school entering first year colleges and universities need to have better aligned skills and strategies taught in the high school classes to be successful in college. Conley (2003b) explains, “student’s need to engage texts critically and identify a theme or idea, and they must be able to identify these before they can move on to any deeper analysis” (p. 17). This is important for teachers to understand so that they can utilize this knowledge and apply to classroom instruction to help prepare students better for college.

There also needs to be a different focus in PreK-12 classes to change how standards and curriculum are delivered (Conley, 2010a). The expectation to improve instruction is to rid the “drill and kill” memorization activities and worksheets and focus more on engaging, challenging curriculum supporting college knowledge and skills (Conley, 2011). To help with this process, the National Governor’s Association (NGA) and the Council of Chief State School Officers (CCSSO) released the *Common Core State Standards* (CCSS, 2012) to help vault public education toward the goal of moving beyond test-prep instruction and more towards the goal of “world-class learning outcomes for all students” (Conley, 2011, p. 17). Through the collaborative effort of experts, school administrators, teachers, and parents, these standards were developed to provide a clear, consistent framework to prepare our children for college and /or the workforce. The standards define the knowledge and skills students acquire in PreK-12 education so that they will graduate high school and be able to succeed in entry-level, credit-bearing academic college courses and workforce training programs. The CCSS standards are:

- Are aligned with college and work expectations
- Are clear, understandable and consistent
- Include rigorous content and application of knowledge through high-order skills;
- Build upon strengths and lessons of current state standards;
- Are informed by other top performing countries, so that all students are prepared to succeed in our global economy and society; and
- Are evidence-based (p. 1)

The purpose of these standards is to increase the rigor and analytical thinking, particularly in the areas of English language arts and mathematics. Improving the key cognitive strategies in the high school curriculum is the critical component for impacting college readiness (Conley 2010a).

Conley's research (2003b, 2005, 2010a, 2010b & 2011) also indicates for students to be prepared for college, high schools need to focus on four dimensions: key cognitive strategies, key content knowledge, academic behaviors, and contextual and awareness skills. Each component interacts with the others to work effectively; unfortunately, the development of key cognitive strategies is the area often overshadowed by high school curriculum and instruction. Furthermore, the fragmented nature of state testing requirements, mainly involving recall and recognition, is not the way to assess for college success. Testing needs to require students to utilize more complex thinking, apply their learning, and demonstrate proficiency in higher forms of cognition (Conley, 2010a). Other suggestions from Conley (2010b) include ensuring students are provided with a core academic program aligned to college readiness standards while also making sure that students have meaningful and challenging work during the senior year. For this to occur in classrooms, state curriculum, professional learning, and teacher instructional techniques and practices need to change. To evaluate this effectiveness, teacher evaluation and appraisal systems need to align to specific outcomes expected for students to obtain necessary key cognitive skills (Darling-Hammond et al., 2012).

Performance Appraisals - Human Resource Management Perspective

Teacher appraisals are used to monitor classroom instruction, determine teacher effectiveness on student achievement, improve job performance, and make human resource (HR) decisions (Toch, 2008). In order to discharge poor performing employees, performance appraisals help to avoid any legal ramifications against termination lawsuits, and provide feedback to improve job performance (Goe, 2007; Fox, 2009; DeNisi & Sonesh, 2011). HR management defines performance appraisals as “a general heading for a variety of activities through which organizations seek to assess employees and develop their competence, enhance performance and distribute reward” (Fletcher, 2001, p. 473).

Specified level of managers, supervisors, or in public schools, the principal, generally conduct yearly evaluations of their employees to evaluate the employee’s contribution to the overall success of the organization (DeNisi & Sonesh, 2011). However, with the staggering results of the number of students who are not prepared to enter college and must take remediation classes in their first year, or have low ACT® or

SAT scores, research indicates there is a gap between teacher evaluations and job performance of teachers who are not preparing our students with the necessary skills to succeed beyond high school (Conley, 2010b). The questions as to why there are so many ineffective teachers who are not being dismissed and why do performance appraisals not accurately represent student success after graduating high school (Chait, 2010; Goldhaber & Anthoney, 2007).

Appraisal Process Issues – Human Resource Management Perspective

When underperforming schools have low levels of teacher dismissal, there is a concern for high school students who have ineffective teachers (Chait, 2010). School districts must hold principals responsible for performance evaluation by instilling a system that ensures evaluators of teachers are rating correctly and are held accountable for accuracy of assigned ratings. Curtis et al., (2005) suggests that raters should justify their employee appraisal scores to their supervisors, yielding more accurate appraisal scores. If school districts want to increase their overall effectiveness as determined by college success of graduating students, then upward accountability is the potential remedy (Curtis, Havey & Rauden, 2005, Conley 2010a).

However, the discrepancy with teacher evaluation scores and underachieving schools can be the result of the principal's fear of how the teacher may respond or react to the overall performance rating, and this may alter the school's cultural climate (Fox, 2009, DeNisi & Sonesh, 2011; Chait, 2010). It is known that human judgment through the use of evaluations can affect job performance, and teachers who are dissatisfied with their appraisal may develop anger, have negative attitudes, and have decreased motivation to improve performance (Becker, & Cropanzano, 2011). However, satisfaction with an evaluation increases job performance, decreases turnover, increases commitment and motivation, and increase the overall level of organizational health (Ellickson & Logsdon, 2002).

Dismissing teachers for poor performance is difficult, and HR specialists indicate that both the teacher and the principal dread the one-to-one meetings, especially if performance ratings are low (Fox, 2009). Teachers expect themselves to be better than average, and if the teacher just meets expectations, he or she can take offense to the rating (Danielson, 2010). In addition, many rating systems contain very subjective categories, and employees feel managers rate certain people higher or lower, not based on actual performance, but rather on their personal feeling toward those individuals (Tziner, Murphy, & Cleveland, 2005).

HR management studies reveal that evaluators have difficult times maintaining neutrality when rating employees. In fact, managers tend to inflate performance ratings to avoid any negative consequences such as resentment towards the manager or loss of promotional chances (Spence & Keeping, 2010). Often, in a school setting, a principal will give a teacher a higher evaluation than deserved, often to avoid excessive paperwork and time needed for improving performance through additional professional development (Kaplan & Owings, 2001). When teachers score below satisfactory on an evaluation, the recommendation is for the principal to explain what areas need improvement and provide necessary resources to help increase job performance (DeNisi & Sonesh 2011). Following the feedback, the evaluator should provide another observation and then celebrate improvements in job performance. It is noteworthy to mention that effective feedback is necessary to determine professional development for low performance appraisals; however it is essential this professional learning aligns to the individual needs of the teacher to improve quality instruction as measured by college readiness (Coggshall, 2012; Conley, 2010b).

It is important that there is a commitment from principals to deliver accurate evaluation scores tied to college preparedness of high school students. Valuing the appraisal process and identifying strongly with the beliefs, goals and philosophy of the school district, provides a greater likelihood the principal will make every effort to do an accurate evaluation (Tziner et al., 2005). By contrast, school districts need to be cautious when using teacher evaluations as a mean for validating rewards, or identifying promotions or salary increases since this too may lead to inaccurate appraisals (Denisi & Sonesh, 2011). In addition, when managers (principals) have high organizational citizenship (defined as cooperative behavior), this contributes to the smooth functioning of an organization, and generally results in more accurate appraisals. Raters who believe in the appraisal system and have confidence in the process will work hard to provide accurate information. Those who report discomfort in monitoring subordinate's performance, provide feedback, and evaluate that performance, are more likely to inflate ratings to avoid negativity (Tziner et al., 2005).

Since the vast majority of research on performance appraisals and evaluation systems are written and developed in the U.S., other countries and cultures find the methods that stem from U.S. research may or may not be effective or applicable in other cultural settings or by employees of different cultural backgrounds (Fletcher, 2001). Research by Fletcher (2001) suggested the extent the behavior of subordinates can be influenced by their superiors and vise versa. In a high power distance culture, superiors and subordinates do not consider themselves as equals, with greater dependency of the latter on the former. In low power distance cultures, there is a greater acceptance of equality and a greater degree of participation and cooperation between those in higher and lower organizational positions. (Fletcher, 2001)

In summary, the theories presented by HR management provide ideas and explanations that perhaps explain inaccurate teacher evaluations and low dismissal rates (Chait, 2010). In addition to providing evidence of the teacher's ability to prepare students for college, measuring teacher effectiveness using academic growth provides quality feedback for teacher improvement (Goldhaber & Anthony, 2007; Danielson, 2010; Donaldson, 2011; Hanushek & Rivkin, 2012). More importantly, our human capital is dependent on students completing college or some post-secondary degree and is critical to the future development of our economy (Fatima, 2009; Hanushek & Rivkin, 2012).

The Principal's Role In Maintaining Effective Teachers

Public schools utilize the school principal or assistant principal as the primary person completing teacher performance appraisals, and it is the principal who is the most important person to ensure excellent teaching occurs (Donaldson, 2011). With the pressure of increasing student achievement and providing all students with effective teachers, principals must have the skills and ability to assess teachers accurately, identifying the quality of instruction, as well as being able to have honest conversations with teachers to help improve instructional practices (Danielson, 2010). Student achievement gains are directly linked to the individual school; but more directly to the individual classroom, and it is the school principal who is charged with managing effective and ineffective teaching (Danielson & McGreal, 2000; Chait, 2010).

The most effective way for the principal to influence student performance is to recruit, hire and retain effective teachers (Beteille, Demetra, & Susanna, 2009; Kersten & Israel, 2005). The disconnect is understanding how and if principals have the skills, ability and knowledge to remove poor performing teachers through the use of the teacher appraisal system. In 2004, the Chicago public school system agreed to a new contract with the Chicago teacher's union, allowing principals the flexibility to dismiss probationary teachers for any reason and without a hearing process. However, the study conducted by Jacob (2011) revealed that "38.8% to 46.2% and 28% to 34% of elementary and secondary principals respectively, including those in some of the worst performing schools in the district, did not dismiss any teacher despite how easy it was to do so under this new policy" (p. 404).

Principals do proceed with caution and tend to select alternative reasons for dismissing teachers. The results from Chicago public school study indicate principals were more likely to dismiss teachers who had poor attendance or history of poor evaluations. However, despite the ease of dismissing teachers with this policy, some principals didn't dismiss teachers at all and in fact some who were dismissed from one campus were rehired in another (Jacob, 2011). Moreover, successful principals who have campuses that have huge academic growth, as measured by student achievement, are more likely to dismiss ineffective teachers and retain effective teachers (Beteille et al., 2009).

Principals have multiple opportunities to ensure quality teaching through the use of teacher evaluations if used appropriately. Teachers surveyed in northwest Florida consistently expressed that having constructive feedback about their strengths or weaknesses with their evaluator is vitally important to professional improvement (Zimmerman & Deckert-Pelton, 2003). The same study also suggests that principals need to be committed to the evaluation process and must demonstrate motivation to be in the classroom evaluating teachers. Since the principal is considered the instructional leader, teachers also expect the principal to have some experience teaching, content-specific knowledge, and a strong pedagogical background (Zimmerman & Deckert-Pelton, 2003).

Performance of students on standardized testing is often a tool to measure hiring decisions made by principals. Principals in Florida admit in a cross-sectional study of staffing practices between elementary schools, that teachers are hired, dismissed and developed based on their overall Florida Comprehensive Assessment Test (FCAT) (Cohen-Vogel, 2011). Findings in the study discovered principals pay particular attention to student test scores to make decisions regarding teaching assignments or reassignments within the schools. By contrast, it is the FCAT achievement data principals have used to identify teacher effectiveness, yet very few teachers are dismissed for poor performance (Weisberg, Sexton, Mulhern, Keeling, 2009; Chait, 2010). Therefore, the sole use of standardized test scores, as an indicator of teacher effectiveness and/or college readiness, is a questionable practice (Conley, 2010b). With large amount of pressure producing high achievement scores on state testing, principals have the responsibility to ensure each classroom has an effective teacher (Chait, 2010).

The majority of teacher evaluation tools used by principals consist of a series of checklists and provide criterion referenced rubrics based on performance indicators claiming to have a positive effect on student achievement (Danielson, 2010; Schmoker, 2006). Principals and assistant principals attend training to learn and understand how to utilize the evaluation tool provided within school districts; however, school administrators often will not do a pre-conference or a pre-observation meeting with teachers, especially if the teacher is tenured (Kersten & Israel, 2005). It is known that in schools with improved student achievement principals work closely with teachers, providing conferencing before and after an observation, as well as ensuring constructive feedback about the use of best instructional strategies (Kaplan & Owings, 2001).

Principals should visit all classrooms at least 10 minutes and as frequently as possible to look for best practices and provide two-way communication. However, 47% of the 192 principals surveyed identified time as an impediment for a successful evaluation process. Kersten and Israel (2005) note “many reported that they were required to evaluate too many teachers each year and that the paperwork demands associated with their district evaluation process were too extensive” (p. 57). Even though this may be a reason why principals claim they are not effectively evaluating teachers, it is the responsibility of the principal to:

- Help weak and incompetent teachers either get better or leave the profession
- Provide intensive support, specific feedback, and professional development when needed

- When teaching improvement does not occur, counsel the employee out of the profession or start nonrenewal proceedings (Kaplan & Owings, 2001, p. 71)

If we are to see substantial gains in preparing students to be successful at college, principals have primary responsibility to ensure quality instruction is occurring and to address poor performance immediately. In addition to maintaining appropriate documentation, the teacher evaluation tool is just one way to ensure quality instruction is occurring in classrooms, and if we are to improve our economic workforce productivity, principals need to be held to higher accountability by ensuring teachers are effectively preparing students for college (Danielson, 2010; Marzano et al., 2011; Hanushek & Rivkin, 2012).

Teacher Evaluations

Education researchers and policy makers analyze instrument tools and evaluation systems as a method to establish the effect of teaching on successful student achievement (Weisberg et al., 2009, Chait, 2010; Darling-Hammond et al., 2012). Since teachers are financed with public tax money, the public demands high-quality teaching; therefore, it is important that annual measures of job performance on teacher effectiveness are measured by student achievement and in particular on the learning gains produced by the teacher (Chait, 2010; Danielson, 2010; Darling-Hammond et al., 2012; Hanushek & Rivkin, 2012).

School administrators use evaluations to provide evidence for necessary professional development, to fix poor quality teaching, and to look for strategies and methods to improve instruction (Danielson 2010; Toch 2008; Gitomer 2011). The teacher evaluation instrument is also a tool used for employment decisions, compensations, and even used to determine if a permanent teacher license should be issued (Danielson, 2010). The reason we need effective teacher evaluation tools is “just as in other professions, every teacher has the responsibility to be involved in career-long quest to improve practice” (Danielson, 2010, p. 37). Human resource management theories suggest performance appraisals should accurately measure job performance and is communicated timely between the evaluator and the teacher to increase job performance (Fox, 2009, DeNisi & Sonesh, 2011, Mitchell, Holtom, &Lee, 2001). However, in the education field, there are multiple concerns and issues with the rating system of teachers since dismissal rates of poor performing teachers are low and capturing effective teaching within evaluations, in terms of preparing students for college, is minimal (Chait, 2010, Conley, 2005).

With concerns of how teachers are not being dismissed for poor performance, the Race to the Top (RTTT) program implemented by the Obama administration, has forced public school districts to find ways to change how teachers are being evaluated This has included the use of student achievement on standardized testing as a measure of teacher effectiveness(Darling-Hammond et al., 2012). The challenge is the inability for administrators to measure effective teaching due to lack of agreement on how to adequately measure effective teaching and how to identify effective teaching strategies (Kane, Taylor, Tyler, & Wooten, 2010). Kane et al. (2010) examined the Teacher Evaluation System (TES) in Cincinnati. Over 90% of teachers were considered proficient or distinguished, yet results on standardized testing indicated otherwise. The fear was that teacher evaluation scores were unrelated to the skills mastered by students. The teachers within the top quartile of performance ratings had no significant correlation to student achievement on standardized tests. The overall results of the TES found, “observation measures of teaching effectiveness are substantively related to student achievement growth, and that some observed teaching practices predict achievement more than others” (Kane et al., 2011, p. 586). The critical component of this research is the ability for a school administrator to evaluate teacher effectiveness accurately as measured by student achievement.

New Teacher Project (2007) also piloted a study that discovered 87% of Chicago’s 600 schools did not see a single “unsatisfactory” teacher rating between 2003 and 2006, especially from the 69 schools declared to be failing by NCLB accountability standards. School administrators gamble with the lives of students when they do not provide teachers the truth about their strengths and weaknesses in the classroom (Conley, 2012). A "widget effect" describes the failure to provide accurate feedback on teacher evaluations and assume teacher effectiveness is the same from teacher to teacher (Weisberg et al., 2009). Large checklists with meaningless criteria in the form of descriptors such as “satisfactory,” “unsatisfactory,” or “outstanding” are provided to teachers, yet no follow up or conversations are conducted to provide improvement efforts (Danielson, 2010). Typical classroom visits consist of an administrator completing a “traditional drive by [principal] a checklist of classroom conditions and teacher behaviors that often don’t focus directly on the quality of instruction” (Toch, 2008, p. 32). Schmoker (2012) describes the wording of criteria and standards used in teacher evaluation rubrics to be ludicrous. In fact, Schmoker (2012) stated, “much of the criteria itself is both misguided and ambiguous - written in [that] thoughtless, tortured prose that continues to mar the education profession" (p. 71). Serious consideration is required to align teacher evaluations to standards of college success of student readiness (Conley, 2010a).

Research also indicates that there is a lack of consistency between evaluators, where some teachers were rated high by one principal, but lower from another (Danielson, 2010). Educational organizations need to develop systems that train and develop analytical judgment that can exercise professional, credible and transparent processes to be used by administrators and teachers in lieu of a broad ranking system (Gitomer, 2011). In addition to inconsistency between evaluations of teachers, Danielson (2010) suggests the need to stop using the same rubric and evaluation tools for beginning teachers as veteran teachers. Evaluators need to have the skills and ability to recognize examples of quality instruction and successful student learning, and then need to be able to engage the teacher in an appropriate level of conversation about their practice (Danielson, 2010; Coggshall, 2012). The feedback needs to be timely and meaningful and should center around the instructional strategies that foster student learning (Curtis et al., 2005; Chait, 2010; Kane et al., 2011).

A powerful catalyst for improving student achievement is the proper use of teacher evaluations that align the teaching evaluation rubric with specific standards tied to effective instructional strategies (Toch, 2008). Marzano et al. (2011) highlights specific researched-based strategies that positively improve student performance when observed in classroom instruction. The use of multiple measures, such as: a portfolio- including lesson plans; videos of the teacher working with students; instructional materials; samples of student work; teachers' written reflective commentary on the videoed lessons; gains of student achievement; and a measurement of student achievement on college entrance exams should be a part of a teacher evaluation (Toch, 2008; Conley, 2012; Hanushek & Rivkin, 2012). All of these measures need to be identified and taught to teachers prior to the evaluation year. Likewise, evaluators need to provide teachers with continuous feedback throughout the year to avoid ambiguity and to confirm the focus is on improving student achievement (Fletcher, 2001; Danielson, 2010; DeNisi & Sonesh, 2011).

The Value-Added Model

The ability to know if students are ready for college depends on the capability to retrieve data captured through methods such as state achievement tests, ACT® or SAT® scores, high school grades, other college entrance exams or quality of teachers (Conley, 2003a, 2010a, 2011; Hanushek & Rivkin, 2011). Educational researchers are interested in studying different levels of individuals and organizational functions, such as students in classrooms who have characteristics such as gender, socioeconomic status and ethnicity, which are independent variables of the classroom (Bryk & Raudenbush, 1988). When individuals are clustered together ("nested") and influence each other, this phenomenon is known as "grouping". To determine how several level of groupings affect each other separately, the statistical model for determining this is known as hierarchical linear modeling (HLM) (Gall, Gall, & Borg, 2007). The Tennessee Value-Added Assessment System (TVAAS), developed by William L. Sanders, is one of the most prominent examples of value-added models (Sanders & Rivers, 1996). This tool is capable of measuring school influences on student success over a period of time and to assess student "progress rather than a percentage of students able to meet an absolute standard" (Ballou, Sanders & Wright, 2004, p.37).

Likewise, the model of HLM responds to criticisms of single level methods and will provide the means to analyze the impact of how a teacher affects college readiness. HLM has become extremely useful in determining how different factors of a student's education can impact learning outcomes. In order to analyze teacher effectiveness on college readiness, HLM can give a statistical formula allowing for change to be measured in addition to assessing multilevel effects that occur simultaneously (Bryk & Raudenbush, 1988). A study conducted by Huang and Moon (2009) used a three-level HLM to analyze 53 schools, 1,544 students, and 154 teachers to determine how different teacher characteristics affected second grade student achievement gains. Huang and Moon (2009) state, "in order to partition out the variance at different levels, several three-level multilevel models were used to analyze the results, isolating the attributable amount of variance to the different levels" (p. 219). This method allows the researcher to separate teacher, student and school variance components to avoid aggregation bias and miss estimation of standard errors (Raudenbush & Bryk, 2002). With new designs of data retrieval systems in public education today, it is becoming much easier for districts and school leadership to look at student achievement gains after certain adjustments to school and student characteristics. This method of study is measured using a combination of HLM statistical formula to determine gains in achievement and is known as the "value-added modeling" (VAM) tool.

Many articles have been published indicating both positive and negative reasons why it is or is not recommended for school districts to use the VAM when determining teacher effectiveness (Baker, Barton, Darling-Hammond, Hartel, Ladd, Linn, Ravitch, Rothstein, Shavelson, & Shepard, 2010; Hanushek & Rivkin, 2012) To begin this comparison, studies have determined that using the VAM as the sole determination of teacher effectiveness may or may not be completely accurate (Green III, Baker, & Oluwole (2012). There are a number of variables negatively affecting the results of the student information and therefore cannot pinpoint the effect completely by that one teacher (Baker et al., 2010; Hammond et al., 2012). The concern is the inability to assure a commonality of all variables presented on VAM tools linking students to teachers (Darling-Hammond et al., 2012). These variables include: student attendance; educated or supportive parents; family resources; student health; family mobility; student assignment to high performing teachers; English Language Proficiency skills; participation in intervention programs; access to libraries or museums; and the funding resources available to schools (Baker et al., 2010; Newton et al. 2010).

Not only is it important to determine if certain variables affect student performance, but it is also important to understand that using achievement test scores as a single predictor is not always a viable way to determine quality teaching (Baker et al., 2010). For example, the Texas Assessment of Knowledge and Skills (TAKS) do not necessarily test students on the content of the current course in which the student is enrolled. Grade 10 math (TAKS) only assesses 45% of Algebra 1 concepts, yet students sitting in Grade 10 are enrolled in geometry classes, learning geometry concepts (Texas Education Agency, 2012a). Therefore, standardized test measures alone are not a reliable indicator of quality teaching nor a predictor of college readiness (Conley, 2010a; Darling-Hammond et al., 2012).

In an alternate point of view, additional evidence indicates the need to be cautious when using VAM as the sole indicator of teacher effectiveness on student achievement. Newton et al. (2010) examined 250 secondary teachers and approximately 3,500 students to reveal how the teacher ranking system of VAM fluctuated, with rankings changing eight deciles depending on the course they taught. These differences in ranking were attributed to student characteristics, socioeconomic status, and parents' educational level. The same study, as cited in Darling-Hammond et al. (2012) conclude using only VAMs to measure teacher effectiveness are inconsistent since "teachers who score in the bottom 20% of rankings in one year, only 20% to 30% had similar ratings the next year, while 25% to 45% of these teachers moved to the top part of the distribution" (p. 9). Braun (2005) clearly indicates the use of VAM results:

> Should not serve as the sole or principle basis for making consequential decisions about teachers. There are many pitfalls to making causal attributions of teacher effectiveness on the basis of the kinds of data available from typical school districts. We still lack sufficient understandings of how seriously the different technical problems threaten the validity of such interpretation. (p.17)

Using only the VAM for teacher evaluations can also develop legal implications if results from the VAM is the only tool used to evaluate teachers. School districts that have policies where 50% of the teacher final evaluation is based on student achievement scores alone may lead to vulnerability to legal challenges (Green III et al., 2012). Green III et al. (2012) conducted a study on the Colorado, Louisiana, and Tennessee's statute on licensed personnel evaluations and determined that, due to the instability of student characteristics, the VAM should not be the sole indicator of teacher performance.

In contrast to reasons why not to use VAM's, Hershberg (2005) notes, "when value-added scores are collected for each classroom and averaged over three years, teachers have rich diagnostic information to improve instruction and administrators have an empirical basis for evaluating teacher effectiveness" (p. 6). The VAM was developed as "an additional measure or as an alternative to reporting student test results as average test scores or percentage of students achieving proficiency on state standards that are required by NCLB" (Mangiante, 2011, p. 44). Therefore, VAM can be used as another method to determine quality teaching. This process presents support to utilize student success on ACT® or SAT® tests to measure growth in college skills and knowledge (Conley, 2012). VAMs can diagnose and measure progress, serve as a guide to instruction, or even be used as a tool for parents to decide what school their child should attend (Sanders & Wright, 2005; ACT® Research and Policy, 2012a). In addition, the use of VAMs can be a multivariate longitudinal data system that provides "wealth of positive diagnostic information available for educational decision makers" (Sanders & Wright, 2005, p. 6). It is highly recommended for school districts to blend a combination of the VAMs with observation-based evaluations on teachers to assess teacher effectiveness especially in low-income or minority schools. Moreover, VAM can provide teachers data about student growth from his or her classroom and as well as aligning professional learning to improve instructional strategies or change curriculum (Mangiante, 2011).

Despite many studies that examine VAM to evaluate teacher effectiveness, Hanushek and Rivkin (2012) provide more recent additional evidence that "value-added measures of teacher effectiveness is not so much about the conceptual approach as it is about the details of the analysis and the interpretation of the results" (p. 132). In the study conducted by Huang and Moon (2009), VAM allowed them to provide student-level control variables to isolate teacher effects, and a variable was selected to construct similarity between teachers and students. In addition to controlled variables, value-added models are designed to adjust for differences in students' prior achievement and socioeconomic status (Sanders & Rivers, 1996).

A similar study by ACT® Research and Policy (2012a), supported by the Bill and Melinda Gates Foundation, established a baseline of academic growth in Algebra 1 and Biology by measuring changes in students' skills and knowledge attributed through the appropriate course. This study used a regression model (form of VAM) to calculate expected scores, and then compared their actual scores to determine if students performed as expected. Results indicated that the use of VAMs assure measurements of growth that can occur for students and VAM scores can be aggregated by teacher to measure teacher effectiveness (ACT® Research and Policy, 2012a). This method of measuring teacher effectiveness through the use of VAM can give schools and school districts information about instructional improvement, especially if a measure of college readiness is incorporated into the data analysis of effective teaching. Some other suggestions of necessary conditions for using VAMs include the following:

1. Ensuring an aligned system; including standards, curriculum, and student assessments within the structure to assess teachers more accurately;
2. To measure teacher effectiveness using student assessments, the assessments should span the same time frame the teacher provided

instruction for that student;

3. Account for other influences and sources of explanatory factors such as prior classroom/teacher (ACT® Research and Policy, 2012b).

VAMs have many advantages such as the ability to measure academic progress of individual students instead of groups of students in the same cohort. It can be used to analyze learning gains of at least one year of growth on all students, not just low performing students, and VAM provide data to improve instruction (Hershberg, 2005). The VAM using hierarchical leveling regression models can provide additional methods to measure teacher effectiveness, if applied correctly, and in combination with other forms of teacher assessment tools (ACT® Research and Policy, 2012a; Mangiante, 2012). Using VAM techniques can also serve as the means to measure teacher effectiveness in terms of preparing students for college and ensuring teachers are providing increased educational knowledge (growth) for students to be successful beyond high school (Conley, 2010a; Sanders & Wright, 2005).

Professional Development and Appraisal System (PDAS)

The passage of the Elementary and Secondary Act of 1965 and later the 1983 publication of *A Nation At Risk: The Imperative for Educational Reform*, called for educational reform across the U.S. (TEA, 2005). Committees were formed to address the quality of teaching in public schools. In 1985 the president of the American Federation of Teachers requested the establishment of national standards for teachers in order to define and assess what good teachers should do in classrooms. These efforts lead to the creation of the National Board for Professional Teaching Standards

(NBPTS) in 1987 (National Board for Professional Teaching Standards, 1991). The NBPTS designed criteria to enhance student achievement through practices and instructional strategies and professional development surrounding these standards. However, when looking at the domains and criteria of the NBPTS or the PDAS, there is no specific mention to any measurement specific to preparing students for college (TEA, 2005; Conley, 2008).

The state of Texas in 1998 passed §21.351 of the Texas Education Code requiring all school districts to select a method of performance appraisal system to evaluate teachers. The Texas commissioner of education developed, designed and recommended all school districts to utilize the Professional Development and Appraisal System (PDAS). This law was designed to focus on increasing student achievement by aligning the teacher appraisal instrument with best practices and over 1000 school districts in the state of Texas adopted this performance measure (Texas Education Agency, 2005).

Using proficiencies described in the *Learner-Centered Schools for Texas: A Vision of Texas Educators* (Texas Education Agency, 2005), teachers are evaluated on 51 criteria over eight domains. The eight PDAS domains are:

Domain I: Active, successful student participation in the learning process

Domain II: Learner-centered instruction

Domain III: Evaluation and feedback on student progress

Domain IV: Management of student discipline, instructional strategies, time and materials

Domain V: Professional communication

Domain VI: Professional development

Domain VII: Compliance with policies, operating procedures and requirements

Domain VIII: Improvement of academic performance of all students on the campus (this data is based on indicators included in the Academic Excellence Indicator System (AEIS)) (Texas Education Agency, 2005, p. 17)

Analyzing the Texas PDAS, Conley (2012) asserts there is no clear process or specific criteria on the evaluation instrument indicating a way to evaluate teachers by providing instructional tools to improve soft skills and academic competencies. Soft skills include a method to observe students successful at working in teams to formulate and solve problems, analyzing information, demonstrating leadership, setting personal goals and demonstrating the ability to take responsibility for own actions. In addition, the PDAS does not provide criteria indicating the teaching of academic competencies known to provide students with successful college skills (TEA, 2005). The academic competencies necessary for college readiness, include students communicating daily in forms of writing, speaking, listening; reading technical documents, interpreting data and solving problems; understanding scientific concepts, principles, rules, laws, and methods; comprehension of social systems and historical frameworks; all with the intention to provide students with the ability to think in terms of postsecondary readiness (Conley, 2012). Comparing these college readiness expectations with the PDAS instrument, the closest criteria that may perhaps align to these recommended skills are within Domain I - criteria 3 and criteria 5 and Domain II - criteria 3. In Domain I - criteria 3 require observable strategies where “student behaviors indicate learning is at a high cognitive level”. Domain I - criteria 5 teachers need to design an environment where “students connect[ing] learning to work and life applications, both within the discipline and with other disciplines.” In Domain II and Domain III - criteria 3, evaluators should also see “instructional strategies that promote critical thinking and problem solving,” (TEA, 2005). When interpreting the rubric of the PDAS instrument, one can conclude there is room for improvement for matching college readiness skills with observable teacher behavior and instructional strategies.

The PDAS instrument requires teachers to be scored and evaluated by at least one certified appraiser, typically a school administrator (principal or assistant principal). The appraiser must hold a supervisor certificate recognized by the state board for education certification (SBEC) and must have attended and completed the required instructional leadership development (ILD) curriculum and certification prior to evaluating a teacher with the PDAS instrument (TEA, 2005). The evaluation of a teacher must be completed once a year and must include at least a minimum of 45 minutes of classroom observations. This amount of classroom observation in combination of focusing on all the PDAS domains has not been researched to determine if there is a correlation between scores and college success of students.

Workforce Productivity (Teacher Index) in Education

Since the General Electric Company’s former CEO, Jack Welch, used a process where he removed the lower 10% of poor performing employees, many companies followed suit such as Microsoft© (www.microsoft.com), Cisco Systems® (www.cisco.com) and Hewlet-Packard© (www.hp.com) (Scullen, Bergey, & Aiman-Smith, 2005). Not only did this policy lead to better employee performance and better leaders within the company, but it also cultivated a climate of high performance and motivation. This process was called the forced distribution rating system (FDRS). There is evidence indicating that workforce productivity improves when a FDRS is used to eliminate the low performing employees who are not performing at the level expected by management (Scullen, Bergey, & Aiman-Smith, 2005). There are many pros and cons associated with this type of practice; however, there is no literature or research conducted with educational researchers indicating how to model similar practices or even the ability to determine teacher productivity. For the purpose of this study, I will refer to teacher productivity as teacher index to perhaps describe capturing teacher effectiveness in terms of preparing students for college.

Even so, Hanushek and Rivkin (2012) measures teacher quality in terms of teacher value-added to determine the implications for policy on ensuring teachers are providing learning gains of students. They also suggest, “eliminating the least effective 5% - 8% of teachers would bring student achievement up by 0.4 standard deviations or higher” (p. 150), and this elimination would increase the added value to the economy by some $70 trillion. Investing in education by providing effective teachers is extremely important, as public education commits to develop and create a viable workforce to support our human capital (Becker, 1993; Conley, 2012; Fatima, 2009; Hanushek & Rivkin, 2012). Data support that graduate and professional degrees provide adults with higher earnings as they become more educated (U.S. Census Bureau, 2012). Moreover, high school student’s benefit from obtaining a post secondary degree by having more global opportunities, and it is the ability to measure this effectiveness that will help reform schools (Hanushek & Rivkin, 2012; Fatima, 2009).

Teacher productivity defined in economic terms is “a measure of the contribution that an ‘input’ makes to the production of some ‘output’ given the level of other inputs used” (Schalock, 1987, p. 59) and in education, this is the amount of learning a student obtains from his or her teacher. Schalock (1987) did a study of determining how to measure teacher productivity by stating it is not something that is easily measured. He utilized a formula called the “index of relative growth” (IRG), which is provided below:

$$IRG = \frac{(\text{post-test}) - (\text{pre-test})}{(100\%) - (\text{pre-test})}$$

Teacher productivity or finding teacher index scores, is simply a measure of learning gains between different time periods using some set of performance standards over the same particular period of time. Since many states are using criterion referenced testing, and curriculum is aligned to these tests, this measures an alignment of student progress toward the standardized outcome of scores on that particular test. For the measuring instrument to be valid, Schalock (1987) suggests that the type of learning must be aligned to type of instruction, and to make sure both conditions are present. Another study using language arts, math and reading SAT® scores, measured teacher productivity in terms of quantity of student achievement gains as the ratio of teacher value-added to a set standard divided by the fixed target score of the 85^{th} percentile (Thum, 2003). This calculation utilizes the California Academic Performance Index to establish the estimate base value for original scores. In addition, the use of a multivariate multilevel calculation is used to model growth and change of the school.

Fatima (2009) studied the effects of professional degrees earned and growth in state workforce productivity to measure the impact of education on the effects of growth in workforce productivity. The “growth was measured as the percent change in the real (inflation-adjusted) gross state product divided by the number of individuals in the labor force – age 16 years and over" (Fatima, 2009, p. 16). Controls were set for initial workforce productivity between set years and for students who migrated to and from different states during the study. In addition, other factors inhibiting workforce productivity included controls for literacy capabilities, high school education, various training and other variety of new technology in various industries. The results of this study indicate and provide consistent data supporting that investing in higher education provides growth in a state's workforce productivity, supporting the human capital theory by providing rationale behind the importance of obtaining effective teachers at the high school level to ensure college success for all students. For the purpose of this study, the Teacher Index model is a formula designed surroundings discussed in this section and will be presented in chapter three.

Literature Summary

Research indicates highly effective teachers have a significant effect on performance of low performing students (Hanushek, 1992; Hanushek & Rivkin, 2010). Students who are assigned to an ineffective teacher will have significantly lower achievement and gains in achievement when compared to those students who have more effective teachers (Hanushek, 1992, Sanders & Rivers, 1996; Schmoker, 2006). However, since school districts tend to use student test scores from achievement tests to determine effectiveness of a teacher, there is not substantial research available indicating that solely using achievement test scores provides a more comprehensive evaluation of the teacher. Furthermore, indicating the use of multiple measures such as including a method to measure college preparedness is recommended (Baker et al., 2010).

To define a highly effective teacher, one must understand the myriad characteristics teachers naturally possess, learn, or develop throughout their teaching career. Teachers with good content knowledge of the subject they teach, in combination with pedagogical instructional techniques to meet the needs of all diverse learners, positively impact student achievement (Looney, 2011). The challenge comes when students are dropping out of college in their first few years because they are not prepared for the rigor of college level classes. College readiness skills are often not embedded in the high school curriculum (Conley, 2003b) and are not assessed accurately in state achievement testing. However, the current practice of teacher evaluation suggests a need to not only change teacher job performance, but to dismiss the teacher when performance, as measured by student achievement on multiple measures, is low (Chait, 2010). Absent or incomplete preparation of our students for sustainability beyond high school is detrimental to our human capital and economic prosperity of the U.S. (Fatima, 2009; Hanushek, 2011). Administrators and educators have the responsibility to prepare students beyond high school by appropriately equipping schools with effective teachers, and evaluating those teachers based on student achievement as determined by meeting college readiness standards (Becker, 1993; Hanushek, 2003; Goe, 2007; Conley, 2010a).

Teacher effectiveness cannot be sufficiently measured by the use of teacher evaluations or the results of student performance based on standardized achievement testing (Chait, 2010; Conley, 2010a; Hanushek & Rivkin, 2012). Human resource management provides theoretical reasons why inaccurate appraisal systems are often overshadowed by education such as not looking at academic growth of students to measure job performance (DeNisi & Sonesh, 2011; Chait, 2010). The literature provides evidence that when evaluation tools are assessed accurately, timely and with constant feedback, there will be an improvement in job performance (Toch, 2008; Fox, 2009). The Jack Welch theory of removing the bottom 10% of the lowest performing employees triggers improved performance and stronger leaders (Scullen, Bergey, & Airman-Smith, 2005). Evaluation tools need to be cognizant of utilizing a wide range of instructional strategies, professional capabilities, pedagogical practices, and professional learning and multiple measures of student data to encompass a thorough evaluation (Marzano et al., 2011). Determining how productive a teacher can be lead to discovering how to measure workforce productivity in terms of teacher effectiveness and impact on college readiness, this area has not been studied (Thum, 2003).

Principals and evaluators need to calibrate their understanding of evaluation tools and ensure they can accurately use the instrument to assess teaching quality observable in a classroom setting (Spence & Keeping, 2010; Tziner et al., 2005; DeNisi & Sonesh, 2011). Furthermore, the observable measures or criteria on the evaluations should accurately reflect student achievement as determined by college readiness (Conley, 2010a). While there are additional considerations, the utilization of value-added models as the sole manner to evaluate teacher effectiveness is not the most conclusive measure of teacher performance. However, these models can be viable options if implemented correctly and accurately (Braun, 2005; Darling-Hammond et al., 2012). When the data from VAMs is analyzed and used correctly to improve instruction, effective teachers can be identified and the process can lead to dramatic affect on the economic growth of our society especially if we can is measured effectiveness by preparing students for college (Conley, 2003a; Hanushek & Rivkin, 2012).

CHAPTER 3

METHODOLOGY

Research Design

This study employs a longitudinal casual comparative research design that examines three different teacher effectiveness models over a three-year time frame. This study determines if current methods of evaluating teachers are indicative of performance on a college entrance exam in a selected Texas school district.

Participants

Teacher Participants

The teacher participants for this study include 115 English Language Arts (ELA) teachers and 172 math teachers from the three high schools examined over a three-year period. As provided in Table 1 and Table 2 the number of participants varied from each high school in both subject areas. School C has the largest number of teachers and school B has the lowest number of teachers. The rationale for including these teachers in the analyses is due to availability of teacher data.

Table 1

ELA Teachers by Gender, Race and Highest Degree Earned

		Campus							
		A		B		C		Total	
		Count	*N* %	Count	N %	Count	*N* %	Count	*N* %
Gender	Female	33	75.0%	21	80.8%	39	86.7%	93	80.9%
	Male	11	25.0%	5	19.2%	6	13.3%	22	19.1%
	Total	44	100.0%	26	100.0%	45	100.0%	115	100.0%
Race	Asian or Pacific Islander	0	.0%	0	.0%	3	6.7%	3	2.6%
	Black, Not of Hispanic Origin	3	6.8%	3	11.5%	0	.0%	6	5.2%
	Hispanic	1	2.3%	0	.0%	1	2.2%	2	1.7%
	White, Not of Hispanic Origin	40	90.9%	23	88.5%	41	91.1%	104	90.4%
	Total	44	100.0%	26	100.0%	45	100.0%	115	100.0%
Highest Degree Earned	Bachelor	28	63.6%	15	57.7%	25	55.6%	68	59.1%
	Master	16	36.4%	11	42.3%	20	44.4%	47	40.9%
	Total	44	100.0%	26	100.0%	45	100.0%	115	100.0%

Table 2

Math Teachers Gender, Race and Highest Degree Earned

		Campus							
		A		B		C		Total	
		Count	*N* %	Count	*N* %	Count	*N* %	Count	*N* %
Gender	Female	38	62.3%	35	79.5%	43	64.2%	116	67.4%
	Male	23	37.7%	9	20.5%	24	35.8%	56	32.6%
	Total	61	100.0%	44	100.0%	67	100.0%	172	100.0%
Race	Asian or Pacific Islander	1	1.6%	4	9.1%	3	4.5%	8	4.7%
	Black, Not of Hispanic Origin	3	4.9%	3	6.8%	7	10.4%	13	7.6%
	Hispanic	3	4.9%	3	6.8%	5	7.5%	11	6.4%
	White, Not of Hispanic Origin	54	88.5%	34	77.3%	52	77.6%	140	81.4%
	Total	61	100.0%	44	100.0%	67	100.0%	172	100.0%
Highest Degree Earned	Bachelor	39	63.9%	27	61.4%	33	49.3%	99	57.6%
	Doctorate	1	1.6%	0	.0%	0	.0%	1	.6%
	Master	21	34.4%	17	38.6%	34	50.7%	72	41.9%
	Total	61	100.0%	44	100.0%	67	100.0%	172	100.0%

Among the participating teachers, 5.2% ELA teachers were identified as African American, 1.7% Hispanic and 90.4% White. Mathematic teacher participants included 7.6% African American, 6.4% Hispanic and 81.4% White. Majority of the participants also hold a bachelor’s degree as displayed in Tables 1 and 2.

Additional teacher demographic variables include total years of experience with the district, total years outside the school district and total years of teaching experience. As indicated in Table 3 the average total years of ELA teaching experience is 13.53 years with a standard deviation of 10.33. Table 4 provides details about the math participants having average total years of experience of 12.5 years and standard deviation of 8.3 years.

Table 3

ELA Teachers Years of Experience

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Years with District	115	25	2	27	7.31	5.252
Years outside District	115	36	0	36	6.22	8.964
Total Years of Experience	115	44	2	46	13.53	10.326
Valid N (listwise)	115					

Table 4

Math Teachers Total Years of Experience

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Years with District	172	21	0	21	7.37	4.749
Years outside District	172	29	0	29	5.13	6.475
Total Years of Experience	172	40	0	40	12.50	8.275
Valid N (listwise)	172					

Student Participants

The overall student participants include 4,861 11th grade students enrolled in English language arts and mathematics in the 2009 through 2012 academic years. The student populations between the three high schools also have slight variance. The overall student population has approximately 12.9% of the students identified as African American (*n* = 628), 14.3% Hispanic (*n* = 697), and 60.9% White (*n* = 2960).

Data in Tables 5, 6 and 7 represent the student demographic variables by individual campuses. All three high schools have an even split of males to females with a distribution of females at 49% in School A (*n* = 777), males at 50.7% (*n* = 798), School B has 49.6% (*n* = 578) females and males at 50.4% (*n* = 587), and School C has 47.9% females (*n* = 1016) and 52.1% males (*n* = 1105). This indicates there is a very close sample population between males and females for this study. In addition Limited English Proficiency (LEP) and special education is displayed by participating campuses. The largest percentage of special education student participants is from School C and School A has the highest number total of LEP students in the three-year period of this study.

The grand total of economically disadvantage students for the three-year period is 23.0% (*N* = 1117). School B contains the largest number of economically disadvantaged students proportionate to total number of students with 30.4% (*n* = 354) and the lowest number of economically disadvantaged students is in School C with only 17.2% (*n* = 365). School A has 25.3% (*n* = 398), which numerically is the highest number of students, but only the second largest proportionate to the total population.

Table 5

School A: Student Demographics

		2010		2011		2012	
		Count	*N* %	Count	*N* %	Count	*N* %
Race	Hispanic	12	2.4%	122	23.7%	141	25.0%
	American Indian/Alaskan Native	2	.4%	3	.6%	2	.4%
	Asian	15	3.0%	15	2.9%	9	1.6%
	Black or African American	73	14.7%	55	10.7%	75	13.3%
	Native Hawaiian/Other Pacific Islander	15	3.0%	5	1.0%	3	.5%
	White	284	57.1%	295	57.3%	324	57.5%
	Two or more	96	19.3%	19	3.7%	9	1.6%
	Total	497	100.0%	515	100.0%	563	100.0 %
Economic Disadvantage	No	363	73.0%	393	76.3%	405	71.9%
	Yes	134	27.0%	122	23.7%	142	25.2%
	Total	497	100.0%	515	100.0%	563	100.0 %
Gender	Female	250	50.3%	258	50.1%	269	47.8%
	Male	247	49.7%	257	49.9%	294	52.2%
	Total	497	100.0%	515	100.0%	563	100.0 %
SPED	No	472	95.0%	489	95.0%	520	92.4%
	Yes	25	5.0%	26	5.0%	43	7.6%
	Total	497	100.0%	515	100.0%	563	100.0 %
LEP	No	477	96.0%	502	97.5%	552	98.0%
	Total	497	100%	515	100%	563	100%

Table 6

School B: Student Demographics

		2010		2011		2012	
		Count	*N* %	Count	*N* %	Count	*N* %
Race	Hispanic	6	1.8%	83	21.7%	105	23.6%
	American Indian/Alaskan Native	3	.9%	4	1.0%	2	.5%
	Asian	11	3.3%	17	4.4%	15	3.4%
	Black or African American	69	20.4%	60	15.7%	79	17.8%
	Native Hawaiian/Other Pacific Islander	5	1.5%	1	.3%	0	.0%
	White	181	53.6%	201	52.5%	231	52.0%
	Two or more	63	18.6%	17	4.4%	12	2.7%
	Total	338	100.0%	383	100.0%	444	100.0%
Economic Disadvantage	No	224	66.3%	288	75.2%	279	62.8%
	Yes	114	33.7%	95	24.8%	145	32.7%
	Total	338	100.0%	383	100.0%	444	100.0%
Gender	Female	168	49.7%	188	49.1%	222	50.0%
	Male	170	50.3%	195	50.9%	222	50.0%
	Total	338	100.0%	383	100.0%	444	100.0%
SPED	No	311	92.0%	353	92.2%	421	94.8%
	Yes	27	8.0%	30	7.8%	23	5.2%
	Total	338	100.0%	383	100.0%	444	100.0%
LEP	No	328	97.0%	380	99.2%	436	98.2%
	Yes	10	3.0%	3	.8%	8	1.8%
	Total	338	100.0%	383	100.0%	444	100.0%

Table 7

School C: Student Demographics

		2010		2011		2012	
		Count	*N* %	Count	*N* %	Count	*N* %
Race	Hispanic	5	.7%	95	14.0%	128	17.7%
	American Indian/Alaskan Native	8	1.1%	4	.6%	5	.7%
	Asian	20	2.8%	21	3.1%	19	2.6%
	Black or African American	72	10.0%	58	8.6%	87	12.0%
	Native Hawaiian/Other Pacific Islander	8	1.1%	3	.4%	0	.0%
	White	494	68.4%	481	71.0%	469	65.0%
	Two or more	115	15.9%	15	2.2%	14	1.9%
	Total	722	100.0%	677	100.0%	722	100.0%
Economic Disadvantage	No	598	82.8%	578	85.4%	556	77.0%
	Yes	124	17.2%	99	14.6%	142	19.7%
	Total	722	100.0%	677	100.0%	722	100.0%
Gender	Female	373	51.7%	308	45.5%	335	46.4%
	Male	349	48.3%	369	54.5%	387	53.6%
	Total	722	100.0%	677	100.0%	722	100.0%
SPED	No	684	94.7%	637	94.1%	678	93.9%
	Yes	38	5.3%	40	5.9%	44	6.1%
	Total	722	100.0%	677	100.0%	722	100.0%
LEP	No	702	97.2%	665	98.2%	709	98.2%
	Yes	20	2.8%	12	1.8%	13	1.8%
	Total	722	100.0%	677	100.0%	722	100.0%

Variables Examined

Dependent Variable

The dependent variables include student performance on the ACT® and SAT® college entrance exams. In the current study, the ACT® scores were measured as a continuous variable with possible scores ranging from 1 to 36 (ACT, 2013) while the SAT® scores are also a continuous variable with scores ranging from 600-1600, using only the verbal/critical reasoning (VCR) and math scores.

The recommended benchmark score on the ACT® that is considered to represent high probability of success in credit-bearing college classes is 18 for the English composition and 22 for math (ACT®, 2013). ACT®, 2013 define success as having a “50% or higher probability of earning a B or higher in the corresponding college course” (p. 3).

The benchmark score of 1045 on the SAT (VCR and math) is associated with 65% of probability of achieving a B- or higher within the first year of college and are more likely to enroll in a four-year university or college (SAT, 2012).

Independent Variables

The independent variables for this study are categorized into two levels for simplicity when discussing hierarchical multilevel analysis. The independent variable includes student achievement measured by the scale score of the Texas Assessment of Knowledge and Skills (TAKS) exit results in math and English language arts during Grade 11. The TAKS exit results are measured as a continuous variable with possible scores ranging from 1,308 to 3,361 and served as a control variable for prior academic knowledge. The TAKS test is a test that assesses the extent to which a student has learned capable to and is capable to apply the defined knowledge and skills at the grade tested (Texas Education Agency, 2012a).

Level 1 independent variables include student demographics. These characteristics comprise of ethnicity, limited English proficiency (LEP), special education (Sped), gender, and economic disadvantage status. Each independent student variable was assigned a numeric value as listed below:

- Economic disadvantage: 1 = yes, 0 = no
- Gender: 1 = female, 0 = male,
- Special Education: 1 = yes, 0 = no
- Limited English Proficiency: 1 = yes, 0 = no,

Level 2 independent variables include specific teacher characteristics. The characteristics analyzed in this study are teacher gender, ethnicity, degree earned, years of experience, years with the district and the composite average evaluation scores from the professional development appraisal system (PDAS). The PDAS instrument is an evaluative rubric tool used by PDAS trained administration (usually the building principal) to rate teachers according to 51 criteria represented within eight domains. Total possible points per Domain include:

- Domain I: 0-24
- Domain II: 0-45
- Domain III: 0-30
- Domain IV: 0-40
- Domain V: 0-30

- Domain VI: 0-16
- Domain VII: 0-15
- Domain VII: 0-50 (TEA, 2005).

The teacher evaluation scores were measured as a categorical variable with scores ranging from 0 to 4, with 0 representing unsatisfactory, 1 representing below expectations, 2 is proficient, and 3 is exceeds expectations. Each domain was scored independently by the PDAS appraiser and rated according to eight PDAS domains listed below:

Domain I: Active, successful student participation in the learning process

Domain II: Learner-centered instruction

Domain III: Evaluation and feedback on student progress

Domain IV: Management of student discipline, instructional strategies, time and materials

Domain V: Professional communication

Domain VI: Professional development

Domain VII: Compliance with policies, operating procedures and requirements

Domain VIII: Improvement of academic performance of all students on the campus (this data is based on indicators included in the academic excellence indicator system (AEIS) (Texas Education Agency, 2005, p. 17).

Teacher demographic variables were measured as a categorical variable. The categorical variables were measured as follows;

- Gender: 1 = female, 0 = male
- Race: African American = 1, Hispanic = 2, White = 3, Other = 4

- Degree earned: Bachelor = 1, Masters = 2, Doctoral = 3

Teacher total years of experience with the district will be measured as a continuous variable.

Procedure

The selected school district granted full permission to access student participant achievement scores from the district's information system for all variables examined in the current study. Initially, the student data files were obtained from the student information systems; eSchool and Schoolnet and organized in an excel spreadsheet. Individual student data was gathered from a 3-year time frame beginning in the 2008-09 and ending in the 2011-12 academic school year for Grade 11 students. Grade 11 students were selected due to data availability of TAKS assessment and participation in in the ACT® and SAT® exam. The exit TAKS test is the assessment each student must pass as a Grade 11 student (junior) in the state of Texas to be eligible for graduation with a recommended high school diploma (Texas Education Agency, 2012a). Additional information will be obtained from the student information system that includes: ethnicity, free or reduced lunch status (socioeconomic status), LEP, Sped status, teacher by course, grade level and year, and college admission tests (ACT® and SAT®) in the areas of math and English. In addition, permission will be granted by the participating school district to access teacher evaluation scores on the PDAS instrument and teacher demographics from the Human Resource office.

Subsequently, the data were entered into my database. Initial calculations included descriptive measures to examine distribution of each variable and to ensure no erroneous entries. Means and standard deviations were calculated for continuous variables while frequencies and percentages were calculated for categorical variables. Then bivariate correlations were calculated to examine the relationships between all the dependent and independent continuous variables. Based on the results of the descriptive and bivariate analysis the data was examined to determine if any transformations are indeed necessary in order to calculate the three models considered in this study.

I then created a table comparing which model had the best predictive value to determine the strength of each evaluation model against the independent variable. Finally, to provide even more statistical support of comparing all three of the teacher evaluations, this study includes a spearman-rho correlation comparison. The three teacher evaluations provided me information on how teachers were scored. This score was then ranked according to the evaluation method used. To measure the strength of each evaluation to which teachers were ranked, I added a spearman-rho correlation to this study. The Spearman-Rho correlation (r_s) "measures the relationship between variables that are placed on an ordinal or categorical scale of measurement" (Creighton, 2007). This correlation produced a rank order correlation coefficient between teacher scores from the PDAS, VAM, and the teacher Index model. Rank comparison tables are provided for each teacher evaluation.

Model 1 – PDAS Model

The college readiness indicators both ACT® and SAT® are regressed upon the PDAS teacher scores among math and ELA teacher. Multiple regression was used, as this is a single level equation with no evidence of any nested effects. The results are regressed using the ACT® and SAT® scores and the average scores which were derived from the eight domains of the PDAS instrument evaluation tool. The PDAS is the evaluation metrics that is a subjective score provided by the principal.

Model 2 – Teacher Index Model

The Teacher index model was calculated using the standard deviation of students' TAKS scores in math and ELA at the classroom level divided by the standard deviation of campus TAKS results within the same subject area multiplied by the coefficient between the two. The results from this ratio became the independent variable where the ACT® or SAT® results were regressed on this score. This model required a single level equation at the campus level. Scores of 1 or greater are indicative of students' variability in performance, meaning if students are more spread out in their performance on the TAKS test compared to the overall campus result leading to the possibility of an ineffective teacher. Scores that are less than one can be an indicator of less variation of student performance and perhaps more effective teaching.

Model 3 – Value-Added Model (VAM)

This study utilized the two-level multilevel model that follows the work of Raudenbush and Bryk (2002). The results of this model were examined to determine if teacher effectiveness was indicative of student performance on the ACT® or SAT® college entrance exams. The dependent variable includes ACT® or SAT® student performance scores. The Level 1 independent variables consist of student characteristics and Level 2 independent variables are teacher characteristics. The following is a graphical representation of the HLM mode using the dependent variables scores as an example:

Level 1 Model

Y = B0 + B1*(ELATAKS) + B2*(GENDER) + B3*(ECD) + B4*(SPED) + B5*(SAFRAM) + B6*(@GT) + R

Level-2 Model

B0 = G00 + G01*(ELAEVAL) + G02*(GENDER) + G03*(DEGREE) + G04*(MINOR) + G05*(EXPERIEN) + U0

B1 = G10

B2 = G20

B3 = G30

B4 = G40

B5 = G50

B6 = G60

Results

The results of the study are reported in tabular format. It was hypothesized that teacher evaluations have no impact on preparing students for college. In fact I predicted that teacher evaluations from the PDAS have a negative predictor and an inaccurate measure as it relates to student college readiness; and Model 3, using a two-level model, will be the closest to prediction of teacher effectiveness on preparing students for college.

CHAPTER 4

INTRODUCTION

The purpose of this study was to compare the use of three teacher evaluations to determine which model is the best predictor to prepare students for college. The three different models include; Model 1 - Professional Development Appraisal System (PDAS), Model 2 - Teacher Index model, and Model 3 - Value Added Model (VAM). The study focused on establishing a connection between each teacher evaluation model and student achievement on SAT® (verbal/critical reasoning and math) and ACT® (math and reading) college entrance exam. The purpose was to determine which method is the best predictor of college readiness and to establish an evaluative tool of determining teacher effectiveness. Evaluation tools are to measure job performance and provide valuable information when making personnel decisions (DeNisi & Sonesh, 2011). The PDAS, Teacher Index model and VAM are all different evaluation tools able to predict teacher effectiveness as measured by college readiness using the measure of student performance on the ACT® and SAT®.

Data was collected on student achievement data from junior level students who took the ACT® reading and math and the SAT® verbal/critical reasoning and math. However, when preliminary descriptive data was completed and we removed erroneous data, the ACT® indicated an extremely low sample size, that forced this study to include only the scores from the SAT®. Regression statistics collected from the three teacher evaluation models is presented below in order to answer the following research questions:

1. Does the Professional Development and Appraisal System (PDAS) instrument predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?
2. Does the teacher index model predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?
3. Does the value-added model (VAM), used to derive a teacher effectiveness score, predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?

This chapter is comprised of three major sections, which include the descriptive data, actually utilized for the study, findings of each model analyzed and then summary of the results of the models (multivariate analysis). Each data table is presented by school year for each English language arts (ELA) teachers and mathematics teachers.

Descriptive Data

This section of the chapter provides the descriptive data based on available research sample and appropriate sample sizes. Therefore, the descriptive measures of each variable are provided.

The descriptive statistics on student performance on the TAKS and SAT® achievement tests is averaged over the three-year of this study and presented in Table 8. Descriptive measures were calculated to ensure there were no erroneous data entries and to study the distribution and shape of the variables. Due to erroneous data discovered from ACT® scores, the ACT® test is not used in this study for final analysis. The ELA TAKS results indicate a range from 1,295 to 3,092; the mean performance was 2,350.02 with a standard deviation of 182.51. The math TAKS results indicate that the performance range was wide with a minimum value of 1,302 and a maximum value of 2,852, mean score of 2,345.71, standard deviation of 214. 96. The mean performance of the SAT verbal/critical reasoning (VCR) indicates a mean score of 537.45, standard deviation of 96.8 over the course of this three-year study and the SAT® math scores ranged from a minimum of 200 to maximum of 800 mean score of 537.45 and standard deviation of 182.512.

Table 8

Average Three-Year Descriptive Statistics on Student TAKS and SAT Scores

	Minimum	Maximum	Mean	Std. Deviation
ELA TAKS	1295	3092	2350.02	182.512
Math TAKS	1302	2852	2345.71	214.964
SAT Verbal/Critical Reasoning	200	800	537.45	96.844
SAT math	210	800	559.20	95.389

Note: Scores are represented as the average over the three-year study

Table 9 provides the descriptive statistics of student performance on the ELA and math TAKS state assessment and the SAT® (SAT- VCR and mathematics) by individual year. There is an increase in the number of students who tested from year to year looking at both the TAKS and SAT® data. In addition, there is an increase of SAT® (VCR) scores from 2010 (*M* = 531.04, *SD* = 98.967) to an increase in 2012 (*M* = 543.28, *SD* = 98.434). The average scores on the SAT® math test did not show as much growth as SAT® (VCR) over the three-year period; 2010 SAT® math (*M* = 557.89, *SD* = 95.490) and 2012 (*M* = 560.48, *SD* = 95.056).

Table 9

Student Performance in ELA TAKS, Math TAKS and SAT by School-Year

		N	Minimum	Maximum	Mean	Std. Deviation
2010	ELA TAKS	1361	1295	3092	2369.20	197.432
	Math TAKS	1347	1302	2839	2325.45	204.376
	SAT Verbal/Critical Reasoning	508	200	800	531.04	98.967
	SAT math	508	270	800	557.89	95.490
2011	ELA TAKS	1396	1354	2854	2344.58	158.503
	Math TAKS	1386	1316	2839	2339.48	211.855
	SAT Verbal/Critical Reasoning	715	270	800	536.28	93.488
	SAT math	715	260	800	558.87	95.760
2012	ELA TAKS	1502	1344	2980	2337.70	187.924
	Math TAKS	1494	1317	2852	2369.76	224.767
	SAT Verbal/Critical Reasoning	702	210	800	543.28	98.439
	SAT math	702	210	800	560.48	95.056

Table 10 provides information about the student characteristics for the independent variable used in this study. After removing erroneous data, this resulted in no limited English-speaking students being utilized in the final analysis of this study for any year, and Hispanic students were removed from the 2010 school year. However, for years 2011 and 2012, Hispanics made up 9.7% (*N* = 166) student participants for this study. For all three years of this study, the African American students accounted for 9.2% (*n* = 157) and White students accounted for 81.2% (*N* = 1391). Addition detailed descriptive statistics identified 8.2% (*N* = 141) economically disadvantage students, 15.6% gifted and talented (*N* = 267) and only 0.7% special education students (*N* = 12). There is also a fairly equal even distribution of males (46.8%, *N* = 803) and females (53.2%, *N* = 911).

Table 10

Descriptive Statistics of Student Characteristics

			School Year			
			2010	2011	2012	Total
Race	Hispanic	Count	0	91	75	166
		% within School Year	.0%	14.0%	11.6%	9.7%
	African American	Count	44	55	58	157
		% within School Year	10.5%	8.5%	9.0%	9.2%
	White	Count	377	502	512	1391
		% within School Year	89.5%	77.5%	79.4%	81.2%
Gender	Male	Count	183	309	311	803
		% within School Year	43.5%	47.7%	48.2%	46.8%
	Female	Count	238	339	334	911
		% within School Year	56.5%	52.3%	51.8%	53.2%
*ED	No	Count	398	592	583	1573
		% within School Year	94.5%	91.4%	90.4%	91.8%
	Yes	Count	23	56	62	141
		% within School Year	5.5%	8.6%	9.6%	8.2%
*SPED	No	Count	420	641	641	1702
		% within School Year	99.8%	98.9%	99.4%	99.3%
	Yes	Count	1	7	4	12
		% within School Year	.2%	1.1%	.6%	.7%
Gifted & Talented	No	Count	352	554	541	1447
		% within School Year	83.6%	85.5%	83.9%	84.4%
	Yes	Count	69	94	104	267
		% within School Year	16.4%	14.5%	16.1%	15.6%
Total		Count	421	648	645	1714

Table 11 presents the descriptive statistics for ELA teacher characteristics. There are 87% females (*N* = 40) representing the ELA teacher subjects in this study. Only 13% of teachers in this study are male (*N* = 6). The highest degree earned by ELA teachers is a bachelors degree, *N* = 26, (56.5%) with 43.5% (*N* = 20) who earned a masters degree or higher. The majority of the ELA teachers in this three-year study are white at 89.1% (*N* = 41) of total sample size with minority teachers representing only 10.9% (*N* = 5).

Table 11

ELA Teacher Gender, Highest Degree Earned, and Race by School Year

		School Year			
		2010	2011	2012	Total
ELA Gender	Male	2	2	2	6
	Male total %	15.3%	12.5%	11.8%	13%
	Female	11	14	15	40
	Female total %	84.6%	87.5%	88.2%	87%
Highest Degree Earned	Bachelor Degree	8	8	10	26
	Bachelor total %	61.5%	50%	58.8%	56.5%
	Graduate Degree	5	8	7	20
	Graduate total %	38.5%	50%	41.2%	43.5%
Race	White	12	14	15	41
	White total %	92.3%	87.5%	88.2&	89.1%
	Minority	1	2	2	5
	Minority total %	7.7%	12.5%	11.8%	10.9%
Total		13	16	17	46

Note: Minority represents African American, Hispanic, Asian, and Other. Graduate degree represents Master and Doctoral degree

The average years of teaching experience for ELA teachers and their average score received on the PDAS evaluation instrument are displayed in Table 12. The average total years of experience for the ELA teachers are 14.4 (*SD* = 11.8). ELA teacher's mean rating score on the PDAS evaluation is 2.45 (*SD* = 11.84).

Table 12

ELA Teacher Total Years of Experience and Average Score on PDAS Evaluation

School Year		ELA PDAS Evaluation	Total Years of Experience
2010	Mean	2.6442	14.3846
	N	21	21
	Std. Deviation	.29247	10.30808
2011	Mean	2.4688	14.6250
	N	18	18
	Std. Deviation	.32755	12.81080
2012	Mean	2.2794	14.1176
	N	25	25
	Std. Deviation	.34098	12.68307
Total	Mean	2.4484	14.3696
	N	64	64
	Std. Deviation	.34918	11.84410

Table 13 is a representation of the math teacher characteristics used in this three-year study. The gender of the math participants included 33.9% males (*N* = 37) and 66.1% females (*N* = 72). The overall number of math participants with a bachelor's degree was 57.8% (*N* = 63) and 42.2% (*N* = 46) of math teachers earned a masters degree or higher. The total number of math teachers used over the span of this three-year study is larger than the total number of ELA teachers (math = 109, ELA = 64). Majority of math teachers are White, *N* = 87 (79.8%) and only 22 (20.2%) were minority, consisting of African American, Hispanic, Asian, and other.

Table 13

Math Teachers Gender, Highest Degree Earned, and Race by School Year

		School Year			
		2010	2011	2012	Total
Math Gender	Male	13	9	15	37
	Male total %	36.1%	25.7%	39.5%	33.9%
	Female	23	26	23	72
	Female total %	63.9%	74.3%	60.5%	66.1%
Highest Degree Earned	Bachelor Degree	21	20	22	63
	Bachelor total %	58.3%	57.1%	57.9%	57.8%
	Graduate Degree	15	15	16	46
	Graduate total %	41.7%	42.9%	42.1%	42.2%
Race	White	30	28	29	87
	White total %	83.3%	80%	76.3%	79.8%
	Minority	6	7	9	22
	Minority total %	16.7%	20%	23.7%	20.2%
Total		36	35	38	109

Note: Minority represents African American, Hispanic, Asian, and Other. Graduate degree represents Master and Doctoral degree

Table 14 displays the evaluation mean score and average total years of experience for the math teachers. The total mean score on the PDAS evaluation is 2.24

(*SD* = 0.289) and total years of experience mean score is 12.44 years (*SD* = 8.73). The highest total years of experience of Math teachers are in 2010, *M* = 13.1 (*SD* = 8.64) and 2012 had the lowest total years of experience with mean score of 11.63 (*SD* = 8.6).

Table 14

Total Years of Experience and Average PDAS Scores for Math Teachers

School Year		MATH PDAS Evaluation	Total Years of Experience
2010	Mean	2.1632	13.1111
	N	36	36
	Std. Deviation	.23111	8.64135
2011	Mean	2.2143	12.6286
	N	35	35
	Std. Deviation	.24172	9.05566
2012	Mean	2.3586	11.6316
	N	38	38
	Std. Deviation	.34530	8.67248
Total	Mean	2.2477	12.4404
	N	109	109
	Std. Deviation	.28917	8.72781

Findings

Model 1 PDAS Model

The data collected from the multiple regression addressed the first research question: How does the Professional Development and Appraisal System (PDAS) instrument, predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?

Separate regressions equations were conducted for the ELA Teachers' PDAS scores and SAT Verbal/Critical Reasoning (VCR) for 2010, 2011, and 2012. Regression equations were also conducted for math teachers' PDAS scores and SAT math in 2010, 2011, and 2012 and are displayed in separate tables by year. In each regression equation, the following student characteristics were entered as control variables in all years: gender, race (African American vs. White), economic disadvantage, special education status, and gifted & talented status. In 2011 and 2012, an additional variable controlled for the effects of race (Hispanic vs. White). All race variables did not have enough students to have a significant output and therefore not used in the calculations.

This model did indicate a significant measure of predicting college readiness as determined by SAT® – verbal critical/reasoning (VCR) and math.

English Language Arts

The SAT was regressed upon the PDAS teacher scores between math and ELA teachers. The eight domains of the PDAS were averaged to provide a composite measure of teacher quality. It was preferable to compute a PDAS composite, rather than enter each of the eight PDAS ratings as separate independent variables, because the PDAS scores were highly correlated. With all eight PDAS scores in the equation, there was very little unique variance shared between a specific PDAS measure and the dependent variable. By contrast, the PDAS was an internally consistent and parsimonious assessment of teacher quality.

The 2010, results of the regressions for SAT VCR is displayed in Table 15. The model summary table provides the *R* square (R^2) explaining the total variability between the dependent variables as explained by average scores on the ELA teacher PDAS evaluation. This is a statistically significant model indicating the significant variables account for 36.3% ($R^2 = 0.363$) variation of dependent variable. The SAT® (VCR) and the PDAS composite were related significantly in 2010, beta = .162, t = 3.997, p < .001. The PDAS has a positive relationship on the dependent variable beta = 61.384

Table 15

Model 1 - Professional Development Appraisal ELA 2010

	Model	Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	*t*	Sig.
1	(Constant)	357.181	41.743		8.557	.000
	African American	-76.416	14.456	-.238	-5.286	.000
	Gender	-.516	7.825	-.003	-.066	.947
	Economic Disadvantage	-18.381	19.212	-.043	-.957	.339
	Special Education	12.626	79.883	.006	.158	.874
	Gifted and Talented	117.248	10.856	.442	10.800	.000
	ELA Teacher Evaluation	61.384	15.356	.162	3.997	.000
R	.602[a]					
R^2	.363					
Adjusted R^2	.354					

Dependent Variable: SAT Verbal/Critical Reasoning

Table 16 displays the *R s*quare information for ELA in 2011. The model explains 34.1% ($R^2 = 0.341$) of the variability in SAT® (VCR) test scores. Specifically, beta = .186, t = 5.766, p < .001. All of the predictor variables during 2011 were significant except student gender. The ELA PDAS evaluation had a positive relationship on SAT® scores, beta = 50.534.

Table 16

Model 1 - Professional Development Appraisal System ELA 2011

	Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	*t*	Sig.
1	(Constant)	399.567	23.078		17.314	.000
	Hispanic	-30.111	9.691	-.111	-3.107	.002
	African American	-42.734	11.619	-.127	-3.678	.000
	Gender	3.172	6.098	.017	.520	.603
	Economic Disadvantage	-39.377	12.244	-.118	-3.216	.001
	Special Education	-87.109	29.230	-.096	-2.980	.003
	Gifted and Talented	119.576	8.650	.447	13.824	.000
	ELA Teacher PDAS Evaluation	50.534	8.765	.186	5.766	.000
R	.584[a]					
R^2	.341					
Adjusted R^2	.339					

Dependent Variable: SAT Verbal/Critical Reasoning

The 2012, results of the regressions for SAT® (VCR) are displayed in Table 17. The model summary table provides the *R* square (R^2) explaining the total variance between the dependent variables as explained by average scores on the ELA teacher PDAS evaluation conducted during the 2012 school year. This is a statistically significant model indicating the significant variables account for 29.5% ($R^2 = 0.295$) variation of dependent variable. The SAT® (VCR) and the PDAS composite were related significantly in 2012, beta = .115, $t = 3.383$, $p < .001$. The PDAS has a positive relationship on the dependent variable beta = 33.738.

Table 17

Model 1 - Professional Development Appraisal ELA 2012

	Model	Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	*t*	Sig.
1	(Constant)	459.221	24.356		18.854	.000
	Hispanic	-25.827	10.975	-.085	-2.353	.019
	African American	-43.540	12.405	-.128	-3.510	.000
	Gender	-2.747	6.526	-.014	-.421	.674
	Economic Disadvantage	-51.578	12.440	-.157	-4.146	.000
	Special Education	-9.100	41.276	-.007	-.220	.826
	Gifted and Talented	105.752	9.020	.401	11.724	.000
	ELA Teacher PDAS Evaluation	33.738	9.973	.115	3.383	.001
R	.543[a]					
R^2	.295					
Adjusted R^2	.287					

Dependent Variable: SAT Verbal/Critical Reading

Comparing the adjusted *R* squared values of the Model 1 for all three years for ELA PDAS, the independent variables decreased each year; 2010 R^2 = 0.363; 2011, R^2 = 0.341 and 2012, R^2 = 0.295

Mathematics

The 2010, results of the regressions for SAT® math and the scores for the math teacher's PDAS evaluation, is displayed in Table 18. The model summary table provides the *R* square (R^2) explaining the total variability between the dependent variables as explained by average scores on the math teacher PDAS evaluation. This is a statistically significant model indicating the significant variables account for 37.8%

(R^2 = 0.378) variation of dependent variable. The SAT (math) and the PDAS math composite were related significantly in 2010, beta = .120, *t* = 2.981, *p* = .003. The PDAS has a positive relationship on the dependent variable beta = 47.059. All predictor variables were significant in this year except economically disadvantage students and special education students.

Table 18

Model 1- Professional Development and Appraisal System Math 2010

		Unstandardized Coefficients		Standardized Coefficients		
	Model	*B*	Std. Error	Beta	*t*	Sig.
1	(Constant)	450.015	34.945		12.878	.000
	African American	-68.876	13.986	-.219	-4.925	.000
	Gender	-15.881	7.536	-.083	-2.107	.036
	Economic Disadvantage	3.004	18.700	.007	.161	.872
	Special Education	14.871	76.232	.008	.195	.845
	Gifted and Talented	127.464	10.370	.500	12.292	.000
	Math Teacher PDAS Evaluation	47.059	15.784	.120	2.981	.003
R	.615[a]					
R^2	.378					
Adjusted R^2	.369					

Dependent Variable: SAT math

The beta weight and adjusted *R s*quared information for the math teachers during 2011 is displayed in Table 19. This table provides information about the variance between predictor variables and scores on the PDAS math regressed upon SAT® (math) indicating $R^2 = 0.365$. The SAT® (math) and the PDAS math composite were related significantly in 2011, beta = .126, $t = 3.966$, $p < .001$. All predictor variables were significant ($p < .001$) for during this model and the PDAS predictor score is beta = 47.017.

Table 19

Model 1 - Professional Development and Appraisal System Math 2011

	Model	Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	*t*	Sig.
1	(Constant)	463.535	27.001		17.167	.000
	Hispanic	-34.990	9.746	-.127	-3.590	.000
	African American	-59.060	11.648	-.171	-5.070	.000
	Gender	-29.171	6.128	-.152	-4.760	.000
	Economic Disadvantage	-29.338	12.272	-.086	-2.391	.017
	Special Education	-112.232	29.408	-.121	-3.816	.000
	Gifted and Talented	117.357	8.719	.431	13.460	.000
	Math Teacher PDAS Evaluation	47.017	11.855	.126	3.966	.000
R	.604[a]					
R^2	.365					
Adjusted R^2	.358					

Dependent Variable: SAT math

Table 20 provides information during the 2012 year on the regression results between the student predictors, including scores on the PDAS evaluation and the SAT® math. This table indicates about 41.1% (R^2 =. 411) of the total variance between SAT® math and the predictors. Specifically, SAT® math scores were related significantly to the PDAS in 2012, beta = .177, t = 5.662, p < .001. The predictor score of the PDAS is 50.392, p < 0.01.

Comparing the adjusted R squared values of the Model 1 for all three years for math PDAS, the independent variables decreased from year; 2010 R^2 = 0.378 and 2011, R^2 = 0.365 then increased during year 2012, R^2 = 0.411.

Table 20

Model 1 - Professional Development Appraisal Math 2012

		Unstandardized Coefficients		Standardized Coefficients		
	Model	*B*	Std. Error	Beta	*t*	Sig.
1	(Constant)	445.513	22.383		19.904	.000
	Hispanic	-16.189	9.576	-.056	-1.691	.091
	African American	-57.721	10.894	-.178	-5.299	.000
	Gender	-27.498	5.723	-.148	-4.805	.000
	Economic Disadvantage	-42.205	10.927	-.134	-3.862	.000
	Special Education	-51.489	36.132	-.044	-1.425	.155
	Gifted and Talented	107.569	7.933	.426	13.559	.000
	Math Teacher PDAS Evaluation	50.392	8.899	.177	5.662	.000
R	.641[a]					
R^2	.411					
Adjusted R^2	.404					

Dependent Variable: SAT math

Model 2: Teacher Index

The data collected from the multiple regression addressed the second research question: How does the teacher index model predict college preparedness among graduating seniors at three selected high schools from a north Texas school district? This model did indicate a significant measure of predicting college readiness as determined by SAT® verbal/critical reasoning (VCR) and math scores and the overall teacher effectiveness score of this model is a better model than Model 1.

English Language Arts

For Model 2, students' SAT® (VCR) scores were regressed onto the teacher index score, as well as the student demographic characteristics: race, gender, economic disadvantage, special education and gifted and talented. As noted earlier, lower scores on the teacher index are interpreted as showing greater teacher effect on preparing students for college.

The results for the regressions for SAT® (VCR) scores are presented in Tables 21, 22 and 23. In 2010, the relationship between SAT® (VCR) scores and the ELA teacher index was not statistically significant, (p = .172). In 2011 and 2012, higher SAT® (VCR) scores were related significantly to higher scores on the ELA teacher index measure. Specifically, higher SAT® (VCR) scores were related significantly to higher ELA teacher index measure scores in 2011, beta = .206, t = 6.377, p < .001, and 2012, beta = .162, t = 4.755, p < .001.

Table 21 represents the relationship between the predictors, including teacher index scores and SAT® scores. During 2010 the ELA teacher index scores are not significant, $p = .172$. Gifted and talented students were the only significant predictor ($p < .001$). Both the PDAS and the teacher index during 2010 had a considerable large amount of insignificant predictor variables indicating issues with sample size.

Table 21

Model 2 - Teacher Index ELA 2010

	Model	Unstandardized Coefficients *B*	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	*t*	Sig.
1	(Constant)	480.513	31.064		15.468	.000
	African American	-78.798	14.699	-.245	-5.361	.000
	Gender	-.460	8.003	-.002	-.057	.954
	Economic Disadvantage	-14.893	19.573	-.034	-.761	.447
	Special Education	-14.239	81.099	-.007	-.176	.861
	Gifted and Talented	124.670	10.876	.470	11.463	.000
	ELA Teacher Index	44.233	32.312	.056	1.369	.172
R	.619[a]					
R^2	.384					
Adjusted R^2	.373					

Dependent Variable: SAT Verbal/Critical Reasoning

Table 22 represents the results of the predictor variables regressed on the SAT® test during 2011. During this year, the predictor variables were all significant ($p < 0.05$) except for gender ($p = .820$). The total variability between SAT scores and the predictor variables are 37.3% ($R^2 = .373$). The teacher index model during this year indicates an unstandardized coefficient weight of beta = 150.693.

Table 22

Model 2 - Teacher Index ELA 2011

	Model	Unstandardized Coefficients *B*	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	*t*	Sig.
1	(Constant)	388.820	22.605		17.200	.000
	Hispanic	-25.462	9.670	-.094	-2.633	.009
	African American	-41.906	11.559	-.124	-3.625	.000
	Gender	1.378	6.067	.007	.227	.820
	Economic Disadvantage	-43.856	12.146	-.131	-3.611	.000
	Special Education	-85.515	29.064	-.094	-2.942	.003
	Gifted and Talented	114.468	8.648	.428	13.236	.000
	ELA Teacher Index	150.693	23.631	.206	6.377	.000
R	.611[a]					
R^2	.373					
Adjusted R^2	.346					

Dependent Variable: SAT Verbal/Critical Reasoning

The *R s*quared value of the Teacher Index model for ELA during 2012 is presented in Table 23. The results from this year have a variance of 36.2% (R^2 = .362). Predictor variables during this school year and the Teacher Index model were all significant (p < .05) except gender (p = .311) and special education (p = .620).

Table 23

Model 2 - Teacher Index ELA 2012

		Unstandardized Coefficients		Standardized Coefficients		
	Model	*B*	Std. Error	Beta	*t*	Sig.
1	(Constant)	423.235	24.903		16.995	.000
	Hispanic	-31.665	10.852	-.105	-2.918	.004
	African American	-44.552	12.309	-.131	-3.619	.000
	Gender	-6.666	6.574	-.035	-1.014	.311
	Economic Disadvantage	-48.004	12.486	-.145	-3.845	.000
	Special Education	23.382	47.178	.017	.496	.620
	Gifted and Talented	101.857	9.013	.389	11.301	.000
	ELA Teacher Index	124.321	26.144	.162	4.755	.000
R	.601[a]					
R^2	.362					
Adjusted R^2	.354					

Dependent Variable: SAT Verbal/Critical Reasoning

Mathematics

The results of the teacher index model for math teachers on average had higher *R* square values than ELA. Table 24 shows that during 2010, teacher index scores were: beta = .115, $t = 2.903$, $p = .004$ and $R^2 = .410$. The predictor variables during this year that were significant were only gender ($p = 0.041$) and gifted and talented ($p < .05$). This table also indicates that for every one point increase in the teacher index score (unstandardized coefficient beta = 81.835), is equivalent to an average increase of 81.835 increase on the SAT® math test.

Table 24

Model 2 - Teacher Index Math 2010

	Model	Unstandardized Coefficients *B*	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	*t*	Sig.
1	(Constant)	487.533	23.366		20.865	.000
	African American	-69.243	14.246	-.219	-4.860	.000
	Gender	-15.518	7.561	-.081	-2.052	.041
	Economic Disadvantage	-1.466	18.746	-.003	-.078	.938
	Special Education	29.547	76.070	.015	.388	.698
	Gifted and Talented	131.064	10.173	.516	12.883	.000
	Math Teacher Index	81.835	28.189	.115	2.903	.004
R	.640[a]					
R^2	.410					
Adjusted R^2	.420					

Dependent Variable: SAT math

During 2011, the SAT® math scores were related significantly with higher math teacher index scores; beta = .235, t = 7.453, p < .001 as shown in Table 25. The R square is .453; explaining 45.3% of the variability in this model is contributed to the predictor variables. All predictor variables during 2011 were significant (p < .05). An increase in the teacher index score on average was equivalent to a large increase on SAT® math scores (beta = 163.731).

Table 25

Model 2 - Teacher Index Math 2011

	Model	Unstandardized Coefficients *B*	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	*t*	Sig.
1	(Constant)	433.423	18.783		23.075	.000
	Hispanic	-34.001	9.469	-.123	-3.591	.000
	African American	-56.643	11.328	-.164	-5.000	.000
	Gender	-28.923	5.956	-.150	-4.857	.000
	Economic Disadvantage	-25.560	12.005	-.074	-2.129	.034
	Special Education	-98.010	28.663	-.106	-3.419	.001
	Gifted and Talented	108.485	8.587	.398	12.633	.000
	Math Teacher Index	163.731	21.967	.235	7.453	.000
R	.673[a]					
R^2	.453					
Adjusted R^2	.447					

Dependent Variable: SAT math

Table 26 shows the variability of SAT® math scores explained by this model ($R^2 = .431$). The results for the regression indicates higher SAT® math scores were associated significantly with higher scores on the math teacher index measure; beta = .271, $t = 8.464$, $p < .001$. During this year two independent variables were not significant; special education ($p = .812$) and Hispanic ($p = .146$). All other independent variables were significant ($p < .05$).

Table 26

Model 2 - Teacher Index Math 2012

	Model	Unstandardized Coefficients *B*	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	*t*	Sig.
1	(Constant)	437.168	16.236		26.926	.000
	Hispanic	-13.637	9.371	-.047	-1.455	.146
	African American	-52.495	10.712	-.161	-4.901	.000
	Gender	-25.445	5.599	-.138	-4.545	.000
	Economic Disadvantage	-43.538	10.854	-.136	-4.011	.000
	Special Education	-9.663	40.612	-.007	-.238	.812
	Gifted and Talented	93.293	8.020	.373	11.632	.000
	Math Teacher Index	167.423	19.781	.271	8.464	.000
R	.657[a]					
R^2	.431					
Adjusted R^2	.425					

Dependent Variable: SAT math

Model 3: Value Added Model

The VAM followed a two-level multilevel model, in which students were nested within teachers. At the student level (Level 1), students' SAT scores were regressed onto student TAKS scores as well as the student demographic characteristics described above. At the teacher level (Level 2), teachers' scores on the PDAS composite measure, as well as teachers' characteristics (race, gender, highest degree earned, and total years of experience). Of particular interest in the present analyses are the effects of the teacher-level PDAS teacher evaluation composite measure in ELA or mathematics.

The results of the HLM analyses for SAT® (VCR) in 2010, 2011, and 2012 are shown in Tables 27, 28 and 29, respectively. The effects for the ELA PDAS Teacher

Evaluation Composite were not significant in 2011. However, higher SAT® (VCR) ELA scores were related significantly with higher teacher scores on the ELA PDAS teacher evaluation composite in 2010, 2012, unstandardized beta = 70.72, *t* = 4.199, *p* = .005, and 2012, unstandardardized beta = 38.64, *t* = 3.240, *p* = .009.

The results of the HLM analyses for SAT® math in 2010, 2011, and 2012 are shown in Tables 30, 31 and 32, respectively. The effects for the teacher level math PDAS teacher evaluation composite measure were barely non-significant in each year.

Table 27

Model 3 - Value Added Model (ELA 2010) Final Estimation of Fixed Effects

			Approx.		
Fixed Effect	Standard Coefficient	Error	T-ratio	d.f.	P-value
Teacher Characteristics					
INTRCPT2, G00	-236.427697	75.56071	-3.129	7	0.018
ELA PDAS Evaluation, G01	70.720369	16.840519	4.199	7	0.005
ELA Gender, G02	10.818311	10.375898	1.043	7	0.332
ELA Highest Degree Earned, G03	1.461053	8.295837	0.176	7	0.866
ELA Race G04	26.908161	23.992444	1.122	7	0.3
ELA Total Years of Experience, G05	-0.339192	0.401355	-0.845	7	0.426
For Student ELA TAKS slope, B1					
INTRCPT2, G10	0.232767	0.025106	9.271	409	0
Student Gender slope, B2					
INTRCPT2, G20	-8.701338	7.276534	-1.196	409	0.233
Student Economic Disadvantage slope, B3					
INTRCPT2, G30	-8.054674	17.745853	-0.454	409	0.65
Special Education slope, B4					
INTRCPT2, G40	38.129698	73.044392	0.522	409	0.601
African American slope, B5					
INTRCPT2, G50	-61.704609	13.45088	-4.587	409	0
Gifted & Talented slope, B6					
INTRCPT2, G60	81.763977	10.702366	7.64	409	0

Table 28

Model 3 - Value Added Model (ELA 2011) Final Estimation of Fixed Effects

Fixed Effect	Standard Coefficient	Error	Approx. T-ratio	d.f.	P-value
Teacher Characteristics					
INTRCPT2, G00	-307.892991	61.557143	-5.002	10	0
ELA PDAS Evaluation, G01	31.707066	19.125343	1.658	10	0.128
ELA Gender, G02	-11.805884	13.310593	-0.887	10	0.396
ELA Highest Degree Earned, G03	11.851909	9.963059	1.19	10	0.262
ELA Race G04	-7.243179	19.034164	-0.381	10	0.711
ELA Total Years of Experience, G05	0.707028	0.515556	1.371	10	0.2
For Student ELA TAKS slope, B1					
INTRCPT2, G10	0.349306	0.024928	14.012	635	0
Student Gender slope, B2					
INTRCPT2, G20	-14.605551	5.429456	-2.69	635	0.008
Student Economic Disadvantage slope, B3					
INTRCPT2, G30	-27.893445	10.654407	-2.618	635	0.009
Student Special Education slope, B4					
INTRCPT2, G40	-28.434664	25.679045	-1.107	635	0.269
Hispanic slope, B5					
INTRCPT2, G50	-17.364217	8.476608	-2.048	635	0.041
African American slope, B6					
INTRCPT2, G60	-29.248505	10.244844	-2.855	635	0.005
Gifted & Talented slope, B7					
INTRCPT2, G70	75.94879	8.638077	8.792	635	0

Table 29

Model 3 - Value Added Model (ELA 2012) Final Estimation of Fixed Effects

Fixed Effect	Standard Coefficient	Error	Approx. T-ratio	d.f.	P-value
Teacher Characteristics					
INTRCPT2, G00	-233.429446	59.815719	-3.902	11	0.003
ELA PDAS Evaluation, G01	38.637873	11.92464	3.24	11	0.009
ELA Gender, G02	-22.067597	10.588628	-2.084	11	0.061
ELA Highest Degree Earned, G03	-23.123318	8.985407	-2.573	11	0.026
ELA Race G04	-0.499743	18.185263	-0.027	11	0.979
ELA Total Years of Experience, G05	1.279388	0.384265	3.329	11	0.007
For Student ELA TAKS slope, B1					
INTRCPT2, G10	0.288876	0.022919	12.604	632	0
Student Gender slope, B2					
INTRCPT2, G20	-18.437843	5.813523	-3.172	632	0.002
Student Economic Disadvantage slope, B3					
INTRCPT2, G30	-26.745229	10.962624	-2.44	632	0.015
Student Special Education slope, B4					
INTRCPT2, G40	-1.117416	36.068132	-0.031	632	0.976
Hispanic slope, B5					
INTRCPT2, G50	-22.899829	9.573411	-2.392	632	0.017
African American slope, B6					
INTRCPT2, G60	-36.62304	10.819987	-3.385	632	0.001
Gifted & Talented slope, B7					
INTRCPT2, G70	77.589972	8.473804	9.156	632	0

Table 30

Model 3 - Value Added Model (Math 2010) Final Estimation of Fixed Effects

Fixed Effect	Standard Coefficient	Error	Approx. T-ratio	d.f.	P-value
Teacher Characteristics					
INTRCPT2, G00	-272.023109	63.049145	-4.314	30	0
Math PDAS Evaluation G01	46.070123	24.532999	1.878	30	0.07
Math Gender, G02	0.561145	12.420058	0.045	30	0.965
Math Highest Degree Earned, G03	6.241668	12.344863	0.506	30	0.616
Math Teacher Race, G04	11.35152	15.917735	0.713	30	0.481
Math Total Experience, G05	-0.526846	0.716757	-0.735	30	0.468
For Student math TAKS slope, B1					
INTRCPT2, G10	0.298533	0.019138	15.599	402	0
Student Gender slope, B2					
INTRCPT2, G20	-9.692105	5.105132	-1.899	402	0.058
Student Economic Disadvantage slope, B3					
INTRCPT2, G30	16.8631	12.987528	1.298	402	0.195
Student Special Education slope, B4					
INTRCPT2, G40	-57.532751	51.57777	-1.115	402	0.266
African American slope, B5					
INTRCPT2, G50	-22.945117	9.851207	-2.329	402	0.02
Gifted & Talented slope, B6					
INTRCPT2, G60	69.275699	8.708346	7.955	402	0

Table 31

Model 3 - Value Added Model (Math 2011) Final Estimation of Fixed Effects

			Approx.		
Fixed Effect	Standard Coefficient	Error	T-ratio	d.f.	P-value
Teacher Characteristics					
INTRCPT2, G00	-265.675774	59.965816	-4.43	29	0
Math PDAS Evaluation G01	43.852751	23.715376	1.849	29	0.074
Math Gender, G02	-0.918567	13.775789	-0.067	29	0.948
Math Highest Degree Earned, G03	-7.364861	13.423194	-0.549	29	0.587
Math Teacher Race, G04	42.865853	17.383516	2.466	29	0.02
Math Total Experience, G05	0.478577	0.675572	0.708	29	0.484
For Student math TAKS slope, B1					
INTRCPT2, G10	0.296499	0.015274	19.412	635	0
Student Gender slope, B2					
INTRCPT2, G20	-22.506622	4.13648	-5.441	635	0
Student Economic Disadvantage slope, B3					
INTRCPT2, G30	-8.917228	8.150114	-1.094	635	0.275
Student Special Education slope, B4					
INTRCPT2, G40	-44.113102	20.470931	-2.155	635	0.031
Hispanic slope, B5					
INTRCPT2, G50	-23.114756	6.474534	-3.57	635	0.001
African American slope, B6					
INTRCPT2, G60	-30.880658	7.874142	-3.922	635	0
Gifted & Talented slope, B7					
INTRCPT2, G70	50.496788	6.713331	7.522	635	0

Table 32

Model 3 - Value Added Model (Math 2012) Final Estimation of Fixed Effects

			Approx.		
Fixed Effect	Standard Coefficient	Error	T-ratio	d.f.	P-value
Teacher Characteristics					
INTRCPT2, G00	-159.730363	51.702211	-3.089	32	0.005
Math PDAS Evaluation G01	27.364368	15.657179	1.748	32	0.09
Math Gender, G02	-4.247451	11.231255	-0.378	32	0.707
Math Highest Degree Earned, G03	0.89919	11.378976	0.079	32	0.938
Math Teacher Race, G04	11.70111	13.59179	0.861	32	0.396
Math Total Experience, G05	0.581105	0.640501	0.907	32	0.371
For Student math TAKS slope, B1					
INTRCPT2, G10	0.26404	0.015857	16.651	632	0
Student Gender slope, B2					
INTRCPT2, G20	-20.664978	4.187824	-4.935	632	0
Student Economic Disadvantage slope, B3					
INTRCPT2, G30	-28.062309	8.102793	-3.463	632	0.001
Student Special Education slope, B4					
INTRCPT2, G40	0.696079	27.203359	0.026	632	0.98
Hispanic slope, B5					
INTRCPT2, G50	-16.72073	7.059432	-2.369	632	0.018
African American slope, B6					
INTRCPT2, G60	-34.382362	8.06824	-4.261	632	0
Gifted & Talented slope, B7					
INTRCPT2, G70	47.378536	6.820462	6.947	632	0

Overall Summary

Comparing the three evaluation models, Table 33 presents each model examined in this study and the *R* squared value for each. The average predictive powers for all three years for ELA teachers indicate all three models are very similar with predictability of effective teaching; Model 1 R^2 = 0.333, Model 2, R^2 = 0.373, and Model 3, R^2 = 0.392. However, there is some difference in predictive power looking at math teachers for each model. Model 1, R^2 = 0.385, Model 2, R^2 = 0.431 and Model 3, R^2 = 0.444. The total overall average of the R-squared values indicate that the Teacher Index model (R^2 = 0.402) and the Value Added model (R^2 = 0.418) provide similar results on determining teacher effectiveness as measured by student scores on SAT.

Table 33

Comparison of Predictive Power All Three Evaluation Models

		Model 1	Model 2	Model 3
Test	Year	*R*-Squared	*R*-Squared	(Level 1 Variance) Explained
ELA	2010	0.363	0.384	0.390
ELA	2011	0.341	0.373	0.399
ELA	2012	0.295	0.362	0.388
ELA Average		0.333	0.373	0.392
Math	2010	0.378	0.410	0.424
Math	2011	0.365	0.453	0.470
Math	2012	0.411	0.431	0.437
Math Average		0.385	0.431	0.444
Total Overall Average		0.359	0.402	0.418

Looking at how each model correlated to one another, a Spearman Rho correlation comparison provides the reader more detail about each model used in this study and how they rank against each other in terms of the teacher evaluation. The

Spearman-Rho correlation (r_s) rank correlation tables are presented. The Spearman-Rho correlation (r_s) "measures the relationship between variables that are placed on an ordinal or categorical scale of measurement" (Creighton, 2007). This correlation produced a rank order correlation coefficient between teacher scores from the PDAS, VAM, and the teacher index model. Rank comparison tables are provided for each teacher evaluation.

Table 34 presents evidence that there is a strong positive correlation between the teacher index model and the value added model (VAM), which has a Spearman's correlation coefficient, r_s is 0.652 (correlation is significant at the 0.01 level, 2-tailed, p = .005). There is no significant correlation between the PDAS and VAM nor the PDAS and the teacher index model. This indicates the teacher index and the VAM have a moderate correlation and similar rankings of teachers.

Table 34

Rank Comparison of All Three Teacher Evaluations – ELA 2011

Spearman-Rho Correlations					
			VAM (HLM)	Teacher Index	PDAS
Spearman's rho	VAM (HLM)	Correlation Coefficient	1.000	.652**	.076
		Sig. (2-tailed)	.	.005	.521
		N	17	17	17
	Teacher Index	Correlation Coefficient	.652**	1.000	.231
		Sig. (2-tailed)	.005	.	.365
		N	17	17	17
	PDAS	Correlation Coefficient	.076	.231	1.00
		Sig. (2-tailed)	.521	.365	.
		N	17	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

Math teacher rank comparison on all three of the teacher evaluations is displayed in Table 35. This data indicates that there is a higher Spearman's correlation between the Teacher Index model and the VAM (r_s = .721, p = .002). The PDAS and the Teacher Index model and the PDAS and the VAM were not significant (p = .625 and p = .843).

Table 35

Rank Comparison of All Three Teacher Evaluations – Math 2011

Spearman-Rho Correlations					
			VAM (HLM)	Teacher Index	PDAS
Spearman's rho	VAM (HLM)	Correlation Coefficient	1.000	.721**	.132
		Sig. (2-tailed)	.	.002	.843
		N	17	17	17
	Teacher Index	Correlation Coefficient	.721**	1.000	.178
		Sig. (2-tailed)	.002	.	.625
		N	17	17	17
	PDAS	Correlation Coefficient	.132	.178	1.000
		Sig. (2-tailed)	.843	.625	.
		N	17	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

A Spearman's-Rho Correlation was also conducted for 2012 on math and ELA teachers to see if there is similarity of results from 2011 rankings. Table 36 does indicate a moderately strong Spearman's correlation coefficient between the Teacher Index and VAM, (r_s = .667 p = .003). There is no significant correlation between the PDAS and Teacher Index (p = .586) nor the PDAS and the VAM (p = .837).

Math teacher rankings for 2012 are represented in Table 37. Similar results of a strong Spearman correlation coefficient between the Teacher Index model and the VAM exists, r_s = .776 (p = .003). The PDAS is not statistically significantly between the VAM and the Teacher Index model.

Table 36

Rank Comparison of All Three Teacher Evaluations – ELA 2012

Spearman-Rho Correlations					
			VAM (HLM)	Teacher Index	PDAS
Spearman's rho	VAM (HLM)	Correlation Coefficient	1.000	.667**	.054
		Sig. (2-tailed)	.	.003	.837
		N	16	16	16
	Teacher Index	Correlation Coefficient	.667**	1.000	.142
		Sig. (2-tailed)	.003	.	.586
		N	16	16	16
	PDAS	Correlation Coefficient	.054	.142	1.000
		Sig. (2-tailed)	.837	.586	.
		N	16	16	16

**. Correlation is significant at the 0.01 level (2-tailed).

Table 37

Rank Comparison of All Three Teacher Evaluations - Math 2012

Spearman-Rho Correlations					
			VAM (HLM)	Teacher Index	PDAS
Spearman's rho	VAM (HLM)	Correlation Coefficient	1.000	.776**	.072
		Sig. (2-tailed)	.	.003	.843
		N	16	16	16
	Teacher Index	Correlation Coefficient	.776**	1.000	.232
		Sig. (2-tailed)	.003	.	.486
		N	16	16	16
	PDAS	Correlation Coefficient	.072	.232	1.000
		Sig. (2-tailed)	.843	.486	.
		N	16	16	16

**. Correlation is significant at the 0.01 level (2-tailed).

CHAPTER 5

DISCUSSION

Introduction

With the No Child Left Behind (NCLB) and the Race To the Top initiative, States were forced to adopt high-stake testing as an essential part of determining student achievement and in particular to determine effective teachers (U.S. Department of Education,2004, 2009, 2012). However, studies prove that using state achievement tests alone are not a good indicator of effective teaching and in particular do not indicate if students are prepared for college (Conley, 2003a, 2003b, 2010a; Goe & Croft, 2009). School districts need to determine if teachers are effective in preparing students for college, and having appropriate evaluations methods capable of measuring teacher effectiveness is critical for public education (Hanushek, 2011; Hanushek & Rivkin, 2012).

Chapter 5 begins with the interpretation of the findings as presented in previous chapters, indication of the best teacher evaluation method analyzed and connections to future research. Findings are presented for each of the evaluation tools measured in this study: Model 1 - Professional Development Appraisal System (PDAS), Model 2 - Teacher Index model, and Model 3 - Value Added Model (VAM). In addition, recommendations regarding the findings are presented. The conclusion connects results of this study to current literature to provide information and recommendations for further research.

Summary of Study

This quantitative study was initiated to do a comparison of three different teacher evaluations to determine if one of the three models presented in the previous chapter was a better predictor of preparing students for college as measured by scores from the SAT® Verbal/Critical Reasoning (VCR) and math college entrance exams. The study provides findings from Model 1 (PDAS) and Model 2 (teacher index) that used a simple multiple regression comparing SAT scores regressed upon the average of the PDAS evaluation tool and the teacher index score. In Model 3, the value-added model (VAM), is a two level hierarchical linear model used to determine a teacher effect score; utilizing students nested within a teachers classroom controlling for all of the student characteristics and teacher characteristics. Student characteristics included race, gender, economic disadvantage status, special education, and gifted and talented. Limited English Speaking students were removed from this study due to low sample numbers once running analysis. Teacher characteristics included race, gender, total years of experience, degree earned, and Texas Assessment of Knowledge and Skills (TAKS) scores on the English Language Arts (ELA) and math test.

Summary of Results and Conclusions

Participant Summary

The analyses focused on three research questions. After screening for outliers, the sample population consisted of a total of 64 English Language Arts (ELA) teachers, 109 math teachers, and a total of 4,835 students. Also noted earlier, there was a low sample size of students over the span of this three-year study that took the ACT® test and therefore the ACT® test was eliminated as an independent variable in this study.

The total number of students, who took the SAT® test vs. the TAKS test over this three-year study, indicated that the school district did not have more than half the students sit to take a college entrance exam. Of the approximate 200 students who did take the SAT® test, the mean score was 537.45 (*SD* = 96.844) on verbal/critical reading and 559.20 on the math (*SD* = 95.389) totaling just above 1,045 as the predictor score to guarantee students a minimum of a B- or better in first year level classes (Conley, 2003a; SAT, 2012a).

Information regarding overall teacher descriptive includes the number of teachers who received similar PDAS evaluation scores from school administrators. For example, the ELA teacher PDAS evaluation, had a mean score of 2.4484 (*SD* = .349). This indicates that all ELA teachers within this school district have been identified as higher than proficiency to exceed expectations. Math PDAS evaluation scores averaged a 2.2477 (*SD* = .28917), also indicating math teachers perform all similar within the district studied. Both mean scores on the PDAS provides information in relationship to mean score to student achievement on SAT®, and that students are barely prepared to achieve a minimum of a B- score. These results support literature that students are not prepared for college and principals are inflating teacher evaluations and consider all teachers to perform equivalently (Chait, 2010; Conley, 2005; Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Donaldson, 2011; Jacob, 2011; Schmoker, 2012)

Model Overview

Results indicated that the best teacher evaluation model for math teachers was

Model 3 – the VAM using a multiple HLM method, followed by Model 2 – the teacher Index and finally the lowest predictor evaluation tool is Model 1 – the PDAS instrument. For ELA teachers, all three models were relatively similar in predicting student scores on the SAT®, however, Model 3 – the VAM was still the best predictor. The predictor values on math indicated that the teacher Index model does have a higher R^2 value than PDAS and is very close to the HLM model. The results from the Spearman-Rho correlations also provided evidence that the VAM using the HLM method has a strong correlation with the teacher Index model, and the PDAS instrument does not correlate to the other two models. The Spearman-Rho correlation produced a rank order correlation coefficient for each model and measures this relationship between all three (Creighton, 2007). Each research question was analyzed by model and by the year of the study.

Research Question 1

Does the Professional Development and Appraisal System (PDAS) instrument, predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?

The regression coefficient results from all three years of the ELA PDAS evaluation are positive in regards to the predictability of student verbal/ critical reasoning (VCR) SAT® scores. The average of the PDAS domains was used, rather than entering each domain separately into the equation. This decision was made since scores on the PDAS domains correlate very highly with one another. A composite average of all the domains is more reliable and internally consistent. The model shown in this study shows that there is unitary, coherent dimension of teacher quality that predicts student achievement as measured by the SAT® (VCR) college exam.

In the 2010-year, the results of this model also indicated there is a negative relationship between African American students on SAT scores and in 2011, African American, Hispanic, economic disadvantage and special education also had a negative relationship. The same student characteristic variables in 2012 did result in a negative relationship except for special education students. During all three years of the study, gender had no significant impact on predictability and gifted and talented (GT) students all had a positive relationship of verbal/critical reasoning SAT scores. Differences in beta coefficients on student characteristics could be contributed to low sample size. In addition, PDAS score differences can also be contributed to different evaluators assigned over the three-year study.

The regression coefficients of the math teacher PDAS evaluations are also positive in regards to the predictability of student math SAT scores. The student variables played a much different result on all three years of the math SAT test than on the verbal reasoning SAT test. For example, in 2010, African American, and gender had a significant negative relationship with the math SAT test, where as economic disadvantage and special education students had no significant affect. Then in 2011, all student variables all had a negative significant except GT students who were positive and with an unstandardized coefficient B = 117.357. Indicating GT students can gain an average 117.357 points on the SAT math. During this year, one can conclude that all characteristics, less GT negatively will impact SAT math scores. The PDAS evaluation scores during 2011 did have a significant positive effect on SAT math scores, meaning, the higher the overall average a math teacher scored on the PDAS evaluation, the higher likelihood, students would perform better on the SAT Math.

The student variables during 2012 indicated a significant ($p < .05$) negative relationship with African American students, gender, economically disadvantage students, and there was no significant affect with Hispanic and special education. No significant affect from Hispanic or special education students could possibly indicate low numbers who either took the test or who performed poorly.

Comparing the PDAS evaluation instrument to the teacher index and the VAM, II included results from a rank correlation using the Spearman-Rho (r_s). The Spearman-Rho is a correlation method to determine if there is any relationship to how teachers were scored using the PDAS in relation to the teacher index model or the VAM model.

The results from this indicate that for both ELA and math teachers, there is no statistically significant correlation between the PDAS instrument and the teacher index or the VAM.

In conclusion, the ELA and math PDAS evaluation is a positive predictor of college readiness as measured by the SAT verbal reasoning and math test, but not as strong as the teacher index or the VAM. This answers Research Question 1, that the PDAS instrument does provide positive correlation and a significant method of predicting college readiness of students, however is the weakest predictor in relation to the other two evaluation methods. This instrument, along with student characteristics, is a statistically significant model accounting for variation of scores on the verbal/critical reasoning and math SAT college entrance exam. The PDAS instrument can be an evaluation tool used to establish the effect of a teacher on successful student achievement on SAT supporting research (Weisberg, Sexton, Mulhern, & Keeling, 2009;

Darling-Hammond et al., 2012). Model 1 also provides an analysis indicative that during all three years on average African American, Hispanic, economically disadvantage, special education students have a negative relationship on the verbal reasoning SAT test and that gender plays into a negative relationship on the math SAT test. Using the PDAS tool, can predict which student group they need to focus on to prepare students for college. This information supports the use of a teacher appraisal method that can successfully determine quality of teaching and provide teachers information on improving instruction (DeNisi & Sonesh, 2011; Danielson, 2010).

Research Question 2

> Does the teacher index model predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?

Model 2 also has a positive relationship of predicting college readiness as measured by the SAT test and an overall higher predictor value than Model 1 (PDAS instrument) and has a strong significant correlation with the VAM instrument but not the PDAS. Even though the year 2010 did not have a significant relationship between the teacher index score and the verbal/critical reasoning SAT test, the overall *R* squared value does indicate at least 30% or above *R* squared values, indicating a dependable model on predicting teacher effect. Overall, the teacher index model was very close in predicting teacher quality, than the PDAS for ELA teachers. However, the teacher index model does show a much higher predictable power than the PDAS for math teachers. Looking at the unstandardized coefficient scores of the teacher index model for all three years for either ELA or math, the beta weights are much higher than PDAS. This indicates that for every increase of teacher performance of an ELA or math, teachers as measured by this model, students on average score 122 points. The PDAS only predicts on average an increase of 48.35 points.

The Teacher Index model is a simple formula that can quickly be used with any assessment to determine quality teaching. Quality teaching provides students with skills and knowledge to be successful in college (Conley, 2010a) and enable students to strive for a higher education (Fatima, 2009). The Teacher Index model is a simple calculation that can be used as an additional evaluation method of determined effective teachers. Using the data from the Spearman-Rho Correlation coefficient tables, the Teacher Index model had the highest correlation with the VAM model looking at the math teacher evaluations in 2011 (r_s = .721) and 2012 (r_s = .776). Research concludes that the Teacher Index model is just as dependable in determining a teacher effectiveness scores as the VAM. Even though literature indicates using student test scores to evaluate teachers is not recommended (Baker, Barton, Darling-Hammond, Hartel, Ladd, Linn, Ravitch, Rothstein, Shavelson, & Shepard, 2010), this study suggests that using a model like the teacher index can provide principals immediate feedback about teacher performance as measured by student achievement.

Research Question 3

> Does the value-added model (VAM) used to derive a teacher effectiveness score, predict college preparedness among graduating seniors at three selected high schools from a north Texas school district?

Model 3 has the best predictive power between all three teacher evaluations used in this study indicating this was the best teacher evaluation used in this study to predict college readiness as determined by performance on the SAT test. The VAM also has a high rank comparison correlation with the teacher index, indicating this convoluted formula can be substituted by using the teacher index simple formula. The VAM evaluation tool has advantages to a researcher such as its ability to provide detail information about the independent variables controlled for this study. This information can be extremely useful when having discussions with teachers since this two-level multiple regression assesses each effect of teacher characteristics and student characteristics (Bryk & Raudenbush, 2002; Haushek & Rivkin, 2012). Model 3 had the highest predictive power of all three models, indicating the best teacher evaluation that is capable of determining teacher effectiveness as measured by college preparedness.

During 2010 and 2011, the ELA standard coefficients indicate that the PDAS evaluation has a positive, significant effect ($p < .005$) on student performing well on the SAT. However, during 2011, this model indicated the PDAS does not have any significant relationship ($p = 0.128$). This could be attributed to the inconsistency of the PDAS score given by the evaluator and student performance on SAT. It could also indicate a different group of teachers during this year.

Teacher race, gender, highest degree earned, and total years of experience, had no significant affect on student performance during all three years of this study. This is supported by literature that teachers who have more education or more years of experience do not necessarily mean they are more effective in the classroom (Clotfelter, Ladd, & Vigdo, 2007; Darling-Hammond & Youngs, 2002). However, in 2012, degrees earned did have a negative effect on student performance but the effect was a negative coefficient. This perhaps can be attributed to an ineffective teacher who had higher than a bachelor degree. Teachers may have the education and certification; however, not knowing good instructional practices can contribute to ineffective teaching (Darling-Hammond, 2006; Kane et al., 2010, Marzano, Frontier & Livingston, 2011).

Reviewing the results of the VAM of the math teachers, the composite PDAS scores for all three years, indicates no significant p-values ($p < .05$). In addition, the ranking of teachers with the PDAS tool has no significant correlation with ranking of teachers using the VAM. This indicates that the VAM is a better evaluation tool when determining teacher effectiveness (Hanushek & Rivkin, 2012). The results of this study also tells the reader, the PDAS evaluations given for math teachers had no affect on how students performed on the SAT. However, using the PDAS evaluation by itself (Model 1), results do show significant positive effect on student achievement. This discrepancy is valuable information for school administrators when deciding on how to evaluate effective teaching. The PDAS is a reliable tool when used correctly (TEA, 2005), however, if you want more detail in terms of controlling for student characteristics, teacher characteristics (including PDAS scores), the VAM does give administrators a great detail of information regarding student achievement and teacher quality as supported by Hanushek & Rivkin (2012). Obtaining a teacher effectiveness score from the VAM is another evaluative tool that provides administrators additional information of how to improve teacher performance (Danielson, 2010).

Limitations of the Study

The data used in this study was obtained from one north Texas school district within a very affluent community. Sample size only included teachers and students with a very different percentage demographics and characteristics than most other districts. The total population of Hispanic students was so low, that this student group could not be used in this study for 2010 results. In addition, this school district did not have enough students who took the ACT® test to do a comparison with the SAT®.

This district also had one large high school where teacher PDAS evaluators are rotated on a yearly basis, making it difficult for the researcher to obtain consistent evaluations on teacher participants from year to year. It was obvious from the standard deviation of mean PDAS scores, evaluators within the district studied, all scored teachers with similar results. When analyzing how each evaluation model correlated against each other using the Spearman-Rho, the researcher noticed a high turnover of teachers each year and difficult to rank the same teachers for all three years.

Recommendations for Practice

This study concentrated on three different teacher evaluation methods to determine if teachers are preparing students for college as measured by scores on the SAT® exam. The results of this study indicate that using the VAM is the highest predictor of determining effective teachers on preparing students for college, followed by the Teacher Index model and then the PDAS instrument. This study also concludes that the Teacher Index model is just as successful and useful as the VAM with strong Spearman-Rho correlations between the two models. This study emphasized the critical need to provide a teacher evaluation tool that aligns with student achievement as measured by a college entrance exam. The research of this study leads to further investigations of quality teaching and determining effective teaching based on college readiness standards. Based on the findings from this research study, the following recommendations should be noted:

1. Conducting a similar study using a much larger sample size of students and teacher participants. Suggestion would be to obtain data from more Texas school districts with different demographics and determine if results are similar to this study.
2. Conduct a similar study using another state evaluation system to determine if that state provides a better teacher evaluation that predicts college readiness of graduating seniors and comparing all methods to one another.
3. The PDAS instrument contains 52 different criterions within the eight domains. I suggest singling out each of the 52 criterions and determining if there is one criterion that is a better predictor of measuring a teacher's effectiveness on preparing students for college. This instrument should be perhaps redesigned by the Texas Education Agency to provide a descriptor that evaluates student success in terms of preparing students for college.
4. School districts and public educators, who conduct yearly evaluations of teacher performance, need to be trained better to identify components within the PDAS instrument. Then a study needs to be conducted to determine if evaluators are consistent with using the instrument.
5. This particular school district did not have enough students who took the ACT® test, therefore a future study of the same models across a larger

sample size using both the SAT® and ACT® could provide a more detailed analysis of all three teacher evaluation methods.

6. With the Value Added Model (VAM) as the best method for predicting college readiness as measured by the SAT® followed by the Teacher Index model, school districts should use both methods to gain information of areas to improve instruction and correlate to student performance on state testing to assess improvement with instruction.
7. Use all three methods and obtain a larger Hispanic and African American student population and determine teacher effects on these two student subgroups.
8. Using same teacher evaluation tools and use student Grade Point Average (GPA) and end of course subject grades to determine if subject grades or GPA is a predictor of college readiness.

Conclusion

According to literature (Goe, 2007; Hairrell, Rupley, Edmonds, Larsen, Simmons, Willson, Byrns, & Vaughn, 2011; Hanushek 1992, 2003, 2011; Hanushek & Rivkin, 2010, 2012) effective teachers do have a direct affect on student achievement. Based on the findings in this study, student achievement as measured by the verbal/critical reasoning and math SAT® college entrance exam is related positively to teacher evaluations and can identify effective teaching. In particular, results from this study indicates using a simple mathematical formula such as the teacher index model, is just as successful to calculate effective teaching as the VAM. Findings also support the literature (Braun, 2005; Chait, 2010; Danielson, 2010) that a teacher evaluation tool can be a form of measuring or accurately reflecting a teacher's effectiveness on preparing students for college. This study concludes that the Texas PDAS instrument can be considered an accurate predictor of determining if teachers are preparing students for college; however, results show that the PDAS does not significantly correlate with the teacher index model or the VAM model. In addition, the PDAS can be used to measure teacher effectiveness and be a reliable instrument for determining job performance to make personal decisions if principals want to know if students are graduating prepared for college (Chait, 2010; Conley, 2008, 2010a, 2012), but is not as reliable as the other two methods. The PDAS instrument has eight different domains where teachers are given a score within criterion that supports the literature of having a way to measure teaching practices as determining teacher effectiveness (Darling-Hammond & Bradsford, 2005; Darling-Hammond et al., 2012). This is reassuring to Texas school administrators when accurately predicting teacher effectiveness. The findings in this study also supports the literature that principals are the key to ensuring excellent teaching is occurring in the class and providing accurate evaluations to measure this and the principal can use the PDAS instrument to manage ineffective teachers who are not preparing students for college (Danielson & McGreal, 2000; Donaldson, 2011).

Literature (Conley, 2003a) also provides research that indicates using standardized testing alone, is not a good indication of determining college readiness. The fact that all three of the models provided a positive relationship of predictability between effective teaching and scores on the SAT® test can be a suggestion for policy makers to include in the evaluation process. In addition, the results from the Spearman-Rho Correlations indicate that the teacher index model is just as successful of predicting teacher effectiveness as the value added model (VAM). This indicates that teachers are ranked similar using the teacher index, as does the VAM. The PDAS teacher evaluation has no statistically significant correlation to either the teacher index or the VAM, indicating teacher evaluation scores are not significantly correlated to either the two models. This supports the conclusion that the teacher index and the VAM are the best teacher evaluation tools used to predict college readiness as measured by student achievement on SAT® scores.

REFERENCES

ACT® Research and Policy (2012a) – *Measuring progress in core high school courses: Insights into value-added measures of teacher effectiveness*. Retrieved 10-26-2012 from http://media.act.org/documents/MeasuringCoreProgress.pdf

ACT® Research and Policy (2012b) *ACT test overview*. Retrieved from http://www.act.org/products/k-12-act-test/

ACT® Research and Policy (2012c) *Profile Report Texas - Graduating Class 2012*. Retrieved from http://www.act.org/newsroom/data/2012/pdf/profile/Texas.pdf

ACT® Research and Policy (2013) *College Readiness Standards*. Retrieved from http://www.act.org/standard/pdf/CRS.pdf

Almendarez, L. (2010). *Human capital theory: Implications for educational development*. Paper Presented at the Belize Country Conference, University of the West Indies. Retrieved from http://www.cavehill.uwi.edu/BNCCde/belize/conference/papers2010/almendarez.html

Babu, S., & Mendro, R. (2003). *Teacher accountability: HLM-based teacher effectiveness indices in the investigation of teacher effects on student achievement in a state assessment program*. Paper presented at the annual meting of the American Education Research Association Illinois. Retrieved from http://www.dallasisd.org/cms/lib/TX01001475/Centricity/Shared/evalacct/research/articles/Babu-Teacher-Accountability-HLM-Based-Teacher-Effectiveness-Indices-2003.pdf

Baker, E.L., Barton, P.E., Darling-Hammond, D., Hartel, E., Ladd, H.F., Linn, R.L., Ravitch, D., Rothstein, R., Shavelson, R.J., & Shepard, L.A. (2010). *Problems with the use of student test score to evaluate teachers* (Briefing Paper 275). Washington, DC: Economic Policy Research. Retrieved from http://www.epi.org/page/-/pdf/bp278.pdf

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

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.M. (2010). Teachers' Mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133-180. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?vid=28&hid=123&sid=3a4c0cf2-fc54-4ce5-8414-3865af399991%40sessionmgr14&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=eric&AN=EJ883791

Becker, G.S. (1993). Nobel lecture: The economic way of behavior. *Journal of Political Economy 101*(3), 385-409. Received from http://libproxy.library.unt.edu:2126/stable/pdfplus/2138769.pdf

Becker, G.S. (1994). *Human Capital: A Theoretical and empirical analysis with special reference to education*. The University of Chicago Press: National Bureau of Economic Research website: http://www.nber.org/books/beck94-1

Becker, W. J. & Cropanzano, R. (2011). Dynamic aspects of voluntary turnover: An integrated approach to curvilinearity in the performance turnover relationship. *Journal of Applied Psychology, 96*(2), 233-246. Retrieved from http://libproxy.library.unt.edu:2110/ehost/pdfviewer/pdfviewer?sid=5c6a436b-e683-462e-ada4-f43e77af9c14%40sessionmgr114&vid=2&hid=109

Beteille, T., Demetra K., & Susanna L. (2009). *Effective Schools: Managing the Recruitment, Development, and Retention of High-Quality Teachers*. CALDER Working Paper 37. Washington, DC: The Urban Institute. Retrieved from http://www.urban.org/UploadedPDF/1001428-effective-schools.pdf

Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Research/pdf/PICVAM.pdf

Bryk, A.S., Raudenbush, S.W. (1988). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. *American Journal of Education 97*(1), 65-108). Retrieved from http://libproxy.library.unt.edu:2077/stable/1084940

Chait, R. (2010). *Removing chronically ineffective teachers, barrier and opportunities*. Center for American Progress. Received from http://www.americanprogress.org/wp-content/uploads/issues/2010/03/pdf/teacher_dismissal.pdf

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2007). How and why do teacher credentials matter for student achievement? National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w12828.pdf?new_window=1

Coggshall, J.G. (2012). *Toward the effective teaching of new college- and career-ready standards: Making professional learning systemic*. National Comprehensive Center for Teacher Quality. Retrieved from http://www.tqsource.org/publications/TowardEffectiveTeaching.pdf

Cohen-Vogel, L. (2011). Staffing to the test: Are today's school personnel practices evidence based? *Educational Evaluation and Policy Analysis*, *33*(4), 483-505. Retrieved from http://libproxy.library.unt.edu:2728/content/33/4/483.full.pdf+html

Common Core State Standards Initiative. (2012). *About the standards*. Retrieved from http://www.corestandards.org/about-the-standards

Conley, D. (2003a). *Mixed messages: What state high school tests communicate about student readiness for college*. Retrieved from University of Oregon, Center for Educational Policy Research website: http://heartland.org/sites/all/modules/custom/heartland_migration/files/pdfs/14093.pdf

Conley, D. (2003b). *Standards for success. Understanding university success: A project of the Association of American Universities and the Pew Charitable Trusts*. Retrieved from University of Oregon, Center for Educational Policy Research website: http://www.epiconline.org/files/pdf/UUS_Complete.pdf

Conley, D (2005). Align high school with college for greater student success. *The Education Digest, 71*(2), 4-12. Retrieved from http://libproxy.library.unt.edu:7125/docview/218180248/fulltextPDF?accountid=7113

Conley, D. (2008). Rethinking college readiness*. New Directions for Higher Education, 2008*(*144)*, 3-13. Retrieved from http://libproxy.library.unt.edu:2179/doi/10.1002/he.321/abstract

Conley, D. (2010a). *College and career ready: helping all students succeed beyond high school*. San Francisco, CA: Jossey-Bass.

Conley, D. (2010b). Eligible & ready for college. *Principal Leadership, 11*(4), 18-22. Retrieved from http://libproxy.library.unt.edu:7125/docview/808594918/fulltextPDF?accountid=7113

Conley, D. (2011). Building on the Common Core State Standards could transform education – if educators translate them into new curriculum and instruction to get student college and career ready. *Educational Leadership 68*(6), 16-20. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?sid=a7b71d33-e5c4-4f99-85f0-a3b3cdcfec23%40sessionmgr114&vid=1&hid=126&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=tfh&AN=58688606

Conley, D. (2012). *A complete definition of college and career readiness*. Retrieved from University of Oregon, Center for Educational Policy Research website: https://www.epiconline.org/readiness/definition.dot

Creighton, T.B. (2007). S*chools and data. The educators guide for using data to improve decision-making*. Thousand Oaks, California: Sage Publications.

Curtis, A. B., Harvey, R.D., & Ravden, D. (2005). Sources of political distortions in performance appraisals: Appraisal purpose and rater accountability. *Group and*

Organization Management, 30(1), 42-60. Retrieved from http://libproxy.library.unt.edu:7125/docview/203352036

Danielson, C. (2010). Evaluations that help teachers learn. *Educational leadership 68*(4), 35-39. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?sid=076e5eb2-0742-4229-92d4-41a7d327c555%40sessionmgr15&vid=1&hid=11&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=tfh&AN=55637399

Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: ASCD.

Darling-Hammond, L., & Youngs, P. (2002). Defining highly qualified teachers: What does scientifically based research. *Educational Researcher*, *31*(9), 13-25. Retrieved from http://libproxy.library.unt.edu:2126/stable/10.2307/3594491?origin=api

Darling-Hammond, L., & Bradsford, J. (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco, CA: Jossey-Bass.

Darling-Hammond, L. (2006). *Powerful teacher education: Lessons from exemplary programs*. San Francisco, CA: Jossey-Bass.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation; Popular modes of evaluating teachers are fraught with inaccuracies and inconsistencies, but the field has identified better approaches. *Phi Delta Kappan, 93*(6), 8-15. Retrieved from http://libproxy.library.unt.edu:2110/ehost/detail?sid=8a9f24f0-6e2b-45fb-aa71-

5935cc1b08cb%40sessionmgr12&vid=1&hid=3&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=tfh&AN=73317375

DeNisi, A. S., & Sonesh, S. (2011). The appraisal and management of performance at work. In, S. Zedeck (Ed.), *APA Handbook of Industrial and Organizational Psychology* (Vol. 2), 255-279. Washington, D.C.: APA.

Donaldson, M. (2011). Principals' approaches to hiring, assigning, evaluating and developing teachers. *Education Digest*, *76*(9), p. 27-32. Retrieved from http://libproxy.library.unt.edu:2063/ehost/pdfviewer/pdfviewer?sid=30a14638-4b15-49ca-bee6-a4c33c371f35%40sessionmgr112&vid=2&hid=106

Ellickson, M.C., & Logsdon, K. (2002). Determinants of job satisfaction of municipal government employees, *Public Personnel Management, 31*(3), 343-358. Retrieved from http://libproxy.library.unt.edu:7125/docview/215944023

Fatima, N. (2009). Investment in graduate and professional degree education: Evidence of state workforce productivity growth. *Florida Journal of Educational Policy 3*(1), 9-35. Retrieved from http://libproxy.library.unt.edu:2063/ehost/pdfviewer/pdfviewer?sid=9bc38cdb-a761-473b-8df2-46c760ed5322%40sessionmgr13&vid=2&hid=15

Ferguson, R.F., & Ladd, H.F. (1996). How and why money matters: An analysis of Alabama schools. In H.F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 265-298). Washington, DC: Brookings Institution Press. Retrieved from http://books.google.com/books?id=UuRDDXgMllwC&printsec=frontcover#v=onepage&q&f=false

Fletcher, C. (2001). Performance appraisal and management: The developing research agenda. *Journal Of Occupational & Organizational Psychology, 74*(4), 473-487. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?vid=5&hid=122&sid=f34d8485-3478-4bac-88d6-332cc7184f3e%40sessionmgr4&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=psyh&AN=2002-10499-005

Fox, A. (2009). Curing what ails performance reviews: Remedies for improving annual appraisals. *HR Magazine 54*(1), 52-56. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?sid=0c2965e6-a272-4edb-abaa-fbc7a2a56ca7%40sessionmgr15&vid=1&hid=8&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=bth&AN=36382764

Gall, M.D., Gall, J.P., Borg, W. (2007) *Educational Research: An introduction* (8th ed.). Boston, MA: Pearson Education Inc.

Gitomer, D. (2011). Road maps for learning and teacher evaluation. *Measurement, 9*(2/3), 146-148. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?sid=3eee328c-cc9f-4796-b54c-ca4b0083822c%40sessionmgr14&vid=1&hid=16&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=a9h&AN=65084677

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Retrieved from National Comprehensive Center for Teacher Quality website: http://www.tqsource.org/publications/LinkBetweenTQandStudentOutcomes.pdf

Goe, L., & Croft, A. (2009). *Methods of evaluating teacher effectiveness*. (Research to practice brief, March 2009). National Comprehensive Center for Teacher Quality website:

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National Board Certification as a signal of effective teaching. *Review of Economics and Statistics 89*(1), 134-150. Retrieved from http://libproxy.library.unt.edu:2104/ehost/pdfviewer/pdfviewer?sid=fb43246f-6dfa-455c-8015-7abdaa450375%40sessionmgr15&vid=2&hid=10

Green III, P.C., Baker, B.D., & Oluwole, J. (2012). The legal and policy implications of value-added teacher assessment policies. *Brigham Young University Education & Law Journal*, *1*, 1-29. Retrieved from http://libproxy.library.unt.edu:2104/ehost/pdfviewer/pdfviewer?sid=c7365284-ee77-483a-b3a9-ca2fcb76aebb%40sessionmgr10&vid=18&hid=15

Hairrell, A., Rupley, W., Edmonds, M., Larsen, R., Simmons, D., Willson, V., Byrns, G., & Vaughn, S. (2011). Examining the Impact of Teacher Quality on Fourth-Grade Students' Comprehension and Content-Area Achievement. *Reading & Writing Quarterly, 27*(3), 239-260. Retrieved from http://untexas.summon.serialssolutions.com/document/show?id=FETCHMERGED-LOGICAL-c1254-4d2920652506cc4e690bba3bb234e782a2e83addd3a30bbdff38f33270da1f811&s.cmd%5B%5D=addFacetValueFilters%28IsScholarly%2Ctrue%29&s.fvf%5B%5D=ContentType%2CJournal+Article%2Cf&s.rf=PublicationDate%2C2011-01-01%3A2012-12-31&t.AuthorCombined=hairrell

Hanover Research (July, 2013). *The ACT and SAT: Success factors and admission trends.* Retrieved from http://www.hanoverresearch.com/wp-content/uploads/2013/07/The-ACT-and-SAT-Success-Factors-and-Admissions-Trends-Featured.pdf

Hanushek, E. (1992). The trade-off between child quantity and quality. *Journal of Political Economy, 100*(1), 84-117. Retrieved from http://libproxy.library.unt.edu:2126/stable/pdfplus/2138807.pdf?acceptTC=true

Hanushek, E. (2003). The failure of input-based schooling policies. *The Economic Journal 113*(485), 68-98. Retrieved from http://libproxy.library.unt.edu:2179/store/10.1111/1468-0297.00099/asset/1468-0297.00099.pdf?v=1&t=h9u4yzy1&s=4a72063aeaedd4d062cfdd03fb2af23e1d8ac0ba

Hanushek, E. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466–479. Retrieved from http://libproxy.library.unt.edu:2127/science/article/pii/S0272775710001718"

Hanushek, E., & Rivkin, S.G. (2010). Generalizations about using value-added measures of teacher quality. The *American Economic Review 100*(2), 267-271. Retrieved from http://libproxy.library.unt.edu:7125/docview/854940149

Hanushek, E., & Rivkin, S.G. (2012). The distribution of teacher quality and implications for policy. *Annual Review of Economics 4*(1), 131-157. Retrieved from http://libproxy.library.unt.edu:3184/doi/abs/10.1146/annurev-economics-080511-111001

Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement.* New York: Rutledge. Retrieved from http://books.google.com/books?hl=en&lr=&id=lh7SZNCabGQC&oi=fnd&pg=PP1&dq=visible+learning+a+synthesis+of+over+800+meta-analysis&ots=dj8JTJUd6C&sig=awNTpAWR-sbwibpz90e_nLrGG2M#v=onepage&q=visible%20learning%20a%20synthesis%20of%20over%20800%20meta-analysis&f=false

Hershberg, T. (2005). Value-added assessment and systemic reform: A response to America's human capital development challenge. *Phi Delta Kappan 87*(4), 276-283. Retrieved from http://libproxy.library.unt.edu:2055/docview/218471150

Hinton, C. & Fischer, K.W. (2010). *Learning from the developmental and biological perspective*, in: H. Dumont, D. Instance & F Benavides (Eds.), Educational Research and Innovation the Nature of Learning: Using Research to Inspire Practice (pp.113-135) Paris: OECD.

Huang, F. & Moon, T. (2009). Is experience the best teacher? A multilevel analysis of teacher characteristics and student achievement in low performing schools. *Educational Assessment, Evaluation & Accountability, 21*(3), p. 209-234. Retrieved from http://libproxy.library.unt.edu:2063/ehost/pdfviewer/pdfviewer?sid=62677d22-2c04-448a-bd05-9f27686ed13f%40sessionmgr110&vid=2&hid=106

Jacob, B. (2011). Do principals fire the worst teachers. *Educational Evaluation and Policy Analysis, 33*(4), p. 403-434. Retrieved from http://libproxy.library.unt.edu:2563/content/33/4/403

Kane, T., Taylor, E., Tyler, J., & Wooten, A. (2010). Identifying effective classroom practices using student data achievement data. *Journal of Human Resources 46*(3), 587-613. Retrieved from http://libproxy.library.unt.edu:2104/ehost/pdfviewer/pdfviewer?sid=9f0877d6-869c-4828-80d0-66d7c36f7f8a%40sessionmgr11&vid=2&hid=28

Kane, T., Taylor, E., Tyler, J., & Wooten, A. (2011). Evaluating teacher effectiveness. *Education Next, 11*(3), 54-60. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?vid=21&hid=118&sid=076e5eb2-0742-4229-92d4-41a7d327c555%40sessionmgr15&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=61203667

Kaplan, L.S., & Owings, W.A. (2001). Teacher quality and student achievement: recommendations for principals. *National Association of Secondary School Principals, 85*(628), 64-73. Retrieved from http://libproxy.library.unt.edu:7125/docview/216032893/fulltextPDF?accountid=7113

Kersten, T.A.,& Israel, M.S. (2005). Teacher evaluation: Principals' insight and suggestions for improvement. *Planning and Changing, 36*(1&2), p. 47-67. Retrieved from http://libproxy.library.unt.edu:2055/docview/218770904

King, S.H., & Watson, A. (2010). Teaching excellence for all our students. *Theory Into Practice, 49*(3), 175-184. Retrieved from http://libproxy.library.unt.edu:2104/ehost/pdfviewer/pdfviewer?sid=58a74139-db2e-46c8-a518-7106e2540f21%40sessionmgr111&vid=2&hid=104

Kline, T. J. B., & Sulsky, L.M. (2009). Measurement and assessment issues in performance appraisal. *Canadian Psychology, 50*(3), 161-171. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?vid=9&hid=122&sid=f34d8485-3478-4bac-88d6-332cc7184f3e%40sessionmgr4&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=pdh&AN=2009-11948-006

Korbin, J.L., Patterson, B.F., Shaw, E.J., Matern, K.D., & Barbuti, S.M. (2008). *Validity of the SAT for predicting first-year college grade point average*. Retrieved from http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2008-5-validity-sat-predicting-first-year-college-grade-point-average.pdf

Looney, J. (2011). Developing high-quality teachers: teacher evaluation for improvement. *European Journal Of Education, 46*(4), 440-455. Retrieved from http://libproxy.library.unt.edu:2179/store/10.1111/j.1465-3435.2011.01492.x/asset/j.1465-3435.2011.01492.x.pdf?v=1&t=h8rfj5p1&s=48f98855d448388cef2ba48b8d2bffa6f81e3f91&systemMessage=Wiley+Online+Library+will+be+disrupted+on+27+October+from+10%3A00-12%3A00+BST+%2805%3A00-07%3A00+EDT%29+for+essential+maintenance

Mangiante, E.M. (2011). Teachers matter: Measures of teacher effectiveness in low-income minority schools. *Educational Assessment, Evaluation & Accountability, 23*(1), 41-63. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?sid=50f2895e-c38a-4ae2-97b2-

57eca973a816%40sessionmgr12&vid=1&hid=25&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=ehh&AN=57721639

Marzano, R.J., Waters, T., (2009) *District leadership that works: Striking the right balance*. Bloomington, ID: Solution Tree.

Marzano, R.J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexander, VA: ASCD.

Mincer, J. (1958). Investment in human capital and personal income distribution. Journal of Political Economy, *66*(4). Retrieved from http://libproxy.library.unt.edu:2126/stable/pdfplus/1827422.pdf?acceptTC=true

Mitchell, T.R., Holtom, B.C., & Lee, T.W. (2001). How to keep your best employees: Developing an effective retention policy. *Academy of Management Executive, 15*(4), 96-107. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?vid=6&hid=8&sid=a325c443-a411-49a2-add2-e4bebe7d5109%40sessionmgr14&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=bth&AN=5897929

Nafukho, F.M., Hairston, N.R., & Brooks, K. (2004) Human capital theory: Implications for human resource development. Human Resource Development International, *7*(4), 545-551. Retrieved from http://libproxy.library.unt.edu:2200/pdf14_16/pdf/2004/5B9/01Dec04/15544716.pdf?T=P&P=AN&K=15544716&S=R&D=ehh&EbscoContent=dGJyMNHX8kSeqLI4zOX0OLCmr0qeprZSsae4SreWxWXS&ContentCustomer=dGJyMPGqsE60rrVKuePfgeyx44Dt6fIA

Namaghi, S. A.O. (2010). A data-driven conceptualization of teacher evaluation. *The Qualitative Report, 15*(6), 1504-1522. Retrieved from http://libproxy.library.unt.edu:7125/docview/854983521/fulltextPDF?accountid=7113

National Board for Professional Teaching Standards, (1991). *Toward high and rigorous standards for the teaching profession (2nd ed.)* Washington, D.C.

National Student Clearing House, (2012). *What we do*. Retrieved from http://www.studentclearinghouse.org/about/what_we_do.php

New Teacher Project. (2007). *Hiring, assignment, and transfer in Chicago Public Schools*. New York: Author. Retrieved from http://tntp.org/assets/documents/TNTPAnalysis-Chicago.pdf?files/TNTPAnalysis-Chicago.pdf

Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Achieves, 18*(23), 1-23. Retrieved from http://libproxy.library.unt.edu:2238/PDFS/EJ913473.pdf

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models; Applications and data analysis methods (2nd ed.)* Thousand Oaks, CA: Sage Publications Inc.

Rice, J. K. (2003). Executive summary and Introduction. In J. K. Rice, Teacher quality: *Understanding the effectiveness of teacher attributes* (pp. v-vii and 1-7). Washington, DC: Economic Policy Institute. Retrieved February 19, 2009, from http://www.epi.org/page/-/old/books/teacher_quality_exec_summary.pdf

Rubin, D.B., Stuart, E.A., & Zanutto, E.L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Statistics, 29*(1), 103-116. Retrieved from http://libproxy.library.unt.edu:2126/stable/10.2307/3701308?origin=api

SAT® (2012a) The SAT report on college & career readiness: 2012. Retrieved from http://media.collegeboard.com/homeOrg/content/pdf/sat-report-college-career-readiness-2012.pdf

SAT® (2012b) SAT report: only 43 percent of 2012 college-bound seniors are ready for college. Retrieved from http://media.collegeboard.com/digitalServices/pdf/research/sat-report-2012-press-release.pdf

SAT (2012c) State profile report: Texas. Retrieved from http://media.collegeboard.com/digitalServices/pdf/research/TX_12_03_03_01.pdf

Sanders, W.L., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Retrieved from University of Tennessee, Value-Added Research and Assessment Center website: http://news.heartland.org/sites/all/modules/custom/heartland_migration/files/pdfs/3048.pdf

Sanders, W.L., & Wright, S.P. (2005). A response to Amrein-Beardsley (2008) "Methodological concerns about education value-added assessment system". Retrieved from http://www.sas.com/resources/asset/Sanders_Wright_response_to_Amrein-Beardsley_4_14_2008.pdf

Schalock, M.D. (1987). What is it? How might it be measured? Mark D. Schalock can it be warranted? *Journal of Teacher Education, 38*(5), p. 59-62. Retrieved from http://libproxy.library.unt.edu:2823/content/38/5/59

Schmoker, M. (2006). *Results now: How we can achieve unprecedented improvements in teaching and learning*. Alexandria, VA: ASCD.

Schmoker, M. (2011). *Focus*. Alexandria, VA: ASCD

Schmoker, M. (2012). The madness of teacher evaluation frameworks. *Phi Delta Kappa 93*(8), 70-71. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?sid=fe411cb8-d248-4487-bde4-dc1f55ecf630%40sessionmgr110&vid=1&hid=123&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=tfh&AN=75122556

Schneider, M., & Yin, L. (2011). *The high cost of low graduation rates: how much does dropping out of college really cost*? Retrieved from American Institute for Research website: http://www.air.org/files/AIR_High_Cost_of_Low_Graduation_Aug2011.pdf

Schochet, P.A., & Chiang, H.S. (2010). *U.S. Department of education: Error rates in measuring teacher and school performance based on student test score gains*. Retrieved from http://ies.ed.gov/ncee/pubs/20104004/pdf/20104004.pdf

Scullen, E.E., Bergey, P.K., & Aiman-Smith, L. (2005). Forced distribution rating systems and the improvement of workforce potential: A baseline simulation. *Personnel Psychology, 58*(1), p. 1-32. Retrieved from http://libproxy.library.unt.edu:2055/docview/220137501

Smith, M. (2012, August 4). Struggling for students' readiness. *The Texas Tribune*. Retrieved from http://www.nytimes.com/2012/08/05/education/most-texas-students-found-not-ready-for-college.html?pagewanted=all&_r=0

Spence, J.R., & Keeping, L.M. (2010). The impact of nonperformance information on ratings of job performance: A policy capturing approach. *Journal of Organizational Behavior, 31*(4), 587-608. DOI: 10.1002/job.648.

Terrell, R.D., & Lindsey, R.B. (2009) *Culturally proficient leadership, the personal journey begins within*. Thousand Oaks, CA: Corwin Press.

Texas Education Agency. (2005). *Professional Development Appraisal System: Teacher Manual*. Retrieved from http://www4.esc13.net/uploads/pdas/docs/PDAS_Teacher_Manual.pdf

Texas Education Agency. (2012a). *Testing and Accountability, TAKS resources*. Retrieved from http://www.tea.state.tx.us/student.assessment/taks

Texas Education Agency. (2012b). *Blueprint for grades 9-11 Mathematics*. Retrieved from http://www.tea.state.tx.us/student.assessment/taks/blueprints/

Thum, M.T. (2003). Measuring progress toward a goal: Estimating teacher productivity using a multivariate multilevel model for value-added analysis. *Sociological Methods Research 32*(2), 153-207. Retrieved from http://libproxy.library.unt.edu:2435/content/32/2/153

Toch, T. (2008). Fixing teacher evaluation. *Education leadership, 66*(2), 32-37. Retrieved from http://libproxy.library.unt.edu:2104/ehost/detail?sid=75ceee81-b36e-45bb-a630-

bd7855faf3d9%40sessionmgr112&vid=1&hid=127&bdata=JnNpdGU9ZWhvc3QtbGl2ZSZzY29wZT1zaXRl#db=tfh&AN=34967537

Tziner, A., Murphy, K. R., & Cleveland, J. N. (2005). Contextual and rater factors affecting rating behavior. *Group and Organization Management, 30*(1), 89-98. Retrieved from http://libproxy.library.unt.edu:2774/content/30/1/89.full.pdf+html

U.S. Census Bureau. (2012). *Educational attainment: 2011 American community survey 1-year estimates*. Retrieved from http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_11_1YR_S1501&prodType=tablehttp://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_11_1YR_S1501&prodType=table

U.S. Department of Education, (2004, March*). New No Child Left Behind Flexibility: Highly qualified teachers fact sheet*. Retrieved from http://www2.ed.gov/nclb/methods/teachers/hqtflexibility.pdf

U.S. Department of Education, (2009, November). *Race to the Top Program Executive Summary*. Retrieved from http://www2.ed.gov/programs/racetothetop/executive-summary.pdf

U.S. Department of Education, (2012, August). *Race to the Top – District Executive Summary*. Retrieved November 4 from http://www2.ed.gov/programs/racetothetop-district/2012-executive-summary.pdf

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect. *Educational Digest, 75*(2), 31-35. Retrieved from http://libproxy.library.unt.edu:7125/docview/218196265/fulltextPDF?accountid=7113

Wong, K.F.E., & Kwong, J.Y.Y. (2007). Effects of rater goals on rating patterns: Evidence from an experimental field study. *Journal of Applied Psychology 92*(2), 577-585. Retrieved from http://libproxy.library.unt.edu:2110/ehost/pdfviewer/pdfviewer?sid=1f6f2739-b919-456a-b173-ddf5382c3307%40sessionmgr104&vid=2&hid=109

Zimmerman, S & Deckert-Pelton, M. (2003) Evaluating the evaluators; Teachers' perceptions of the principal's role in professional evaluation. *NASSP Bulletin 87* (636), p. 28-37. Retrieved from http://libproxy.library.unt.edu:2879/content/87/636/28.full.pdf+html