

A CONFIRMATORY FACTOR ANALYSIS OF A MEASURE OF PRESERVICE
TEACHER KNOWLEDGE, SKILLS, AND DISPOSITIONS

by

Antoinette Crawford Ellison

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Education in
Educational Leadership

Charlotte

2009

Approved by:

Dr. Claudia P. Flowers

Dr. Daniel S. Dupre

Dr. Victoria P. Jaus

Dr. Do-Hong Kim

©2009
Antoinette Crawford Ellison
ALL RIGHTS RESERVED

ABSTRACT

ANTOINETTE CRAWFORD ELLISON. A confirmatory factor analysis of a measure of preservice teacher knowledge, skills and dispositions. (Under the direction of DR. CLAUDIA P. FLOWERS)

Recent literature on school improvement indicates that teacher quality has a large impact on student learning (Aaronson, Barrow, & Sander, 2007; Rivkin, Hanushek & Kane, 2005; Hanushek, Kane, O'Brien, & Rivkin, (2005); Nye, Konstantopoulos, Hedges, 2004; Rockoff 2004). The most effective way to create quality teachers is to identify characteristics of teacher quality, measure those characteristics and provide feedback that promotes professional development. Valid and reliable outcomes from instruments that are based on effective teaching standards are needed to provide feedback to teachers. The Student Teacher Assessment Rubric (STAR) is an evaluation tool designed specifically to be used with student teachers. The purpose of this study was to examine the underlying structure of the STAR using confirmatory factor analysis (CFA).

The data were divided into two data sets, with one data set used to examine the fit to the 10-factor model and the second data set used to validate the model. Due to high correlation coefficients among the 10 latent variables, the model specification was changed to a one-factor model. The fit statistics from the CFA for the one-factor model suggested an adequate fit but the number of modifications needed to improve the fit suggested some problems. Implications for measuring complex knowledge and skills needed for effective teaching are discussed.

ACKNOWLEDGEMENTS

I would like to express thanks to the members of my dissertation committee, Dr. Claudia P. Flowers, Dr. Daniel S. Dupre, Dr. Victoria P. Jaus, and Dr. Do-Hong Kim. I would especially like to thank Dr. Claudia Flowers for persevering with me as my advisor throughout the time it took me to complete this research and write the dissertation. Her words of encouragement, quiet urgings and careful reading of all of my writing will never be forgotten.

I would like to thank my husband, William Ellison for his patience, support and unwavering faith in me. In my moments of doubt, he was a constant cheerleader and supporter. I greatly appreciate you for believing in me.

I would like to thank my parents, Willie and Bertha Crawford, for always stressing the importance of education and pushing me to be my best self. Thank you for your many prayers and calming conversations when the task seemed impossible.

I would like to thank God who gave me the strength and perseverance to continue through the challenging times. I acknowledge that without God, none of this would be possible.

TABLE OF CONTENTS

LIST OF TABLES	viii
CHAPTER I: NATURE AND SIGNIFICANCE OF THE STUDY	1
Statement of the Problem	6
Purpose of the Study	7
Research Questions	7
Delimitations of the Study	8
Limitations of the Study	8
Assumptions	8
Operational Definitions	9
Significance of the Study	10
Design and Overview of the Study	11
Summary	12
CHAPTER II: REVIEW OF LITERATURE	14
National Board for Professional Teaching Standards	16
National Council for Accreditation of Teacher Education	17
American Association of Colleges for Teacher Education	21
Interstate New Teacher Assessment and Support Consortium Standards	21
Measures of General Academic Ability	26
Knowledge of Teaching and Learning	33
Advanced Degree	36
Teaching Experience	38

Type of Certification	40
Teacher Performance Evaluation	50
Preservice Teacher Performance Evaluation	55
Summary	59
CHAPTER III: METHODOLOGY AND PROCEDURES	60
Description of Population and Research Setting	62
Instrumentation	64
Research Methods	65
Procedures	65
Data Collection	65
Model Specification	66
Model Identification	68
Model Estimation	68
Model Testing	68
Model Modifications	69
Summary	69
CHAPTER IV: PRESENTATION OF THE DATA	70
Description of Sample	70
Data Screening	72
CFA Results	74
Model Specification	74
Model Identification	75
Model Estimation	75

Model Testing	75
Model Modifications	76
Model Validation	76
Summary	78
CHAPTER V: FINDINGS, IMPLICATIONS, FUTURE RESEARCH	79
Key Findings	80
Implications	80
Future Research	83
Summary	84
REFERENCES	85

LIST OF TABLES

TABLE 1: INTASC Teaching Standards	23
TABLE 2: Latent and Observed Variables	67
TABLE 3: Description of Participants Degree Program and Area of Licensure	72
TABLE 4: Means and Standard Deviations for STAR Ratings	73
TABLE 5: Latent Variables Correlation Matrix	76
TABLE 6: Path Coefficients for Data Set 1	77
TABLE 7: Path Coefficients for Data Set 2	78

CHAPTER I: NATURE AND SIGNIFICANCE OF THE STUDY

The most critical challenge facing public schools across the nation is placing a highly qualified teacher in every classroom. This challenge is becoming increasingly difficult to meet in light of the most dramatic teacher shortage in the history of public schools.

Over 50 years of research has consistently documented that teacher quality is the key factor in student learning. Recent literature on teacher quality overwhelmingly indicates that differences in teacher quality have a large impact on student learning (Aaronson, Barrow, & Sander, 2007; Rivkin, Hanushek & Kane, 2005; Hanushek, Kane, O'Brien, & Rivkin, (2005); Nye, Konstantopoulos, Hedges, 2004; Rockoff 2004). Teachers are the most critical school-based component in the academic progress of students. Having a highly effective teacher for multiple years can mitigate the effect of poverty and low parental levels of education (Fuller, Carpenter & Fuller, 2008). Furthermore, research shows that teacher quality more heavily influences differences in student performance than race, class, or individual school and that teacher quality has a substantial impact on the student achievement of disadvantaged students (Nye et al., 2004). Teacher quality also stands out in research for its potential to close the gap in academic achievement between students from traditionally poor, non-white, and/or urban backgrounds and their middle class, white, suburban peers (Haycock, 2001). Achievement gains as a result of having a quality teacher can be almost three times as large for African American students as for white students, even

when comparing students with the same prior school achievement (Kaplan & Owings, 2001).

Sanders and Rivers (2001) documented the cumulative effects of teacher quality. They compiled standardized test scores in a variety of academic fields from approximately six million students since 1992. In addition to assessment data, they collected data on a wide range of student variables. This longitudinal approach allowed them to assess student progress from year to year. This study is considered to be a valuable piece of research that supports several key factors. Sanders and Rivers (2001) were able to measure the effect of teachers on students while separating ethnic, socioeconomic and parental influences. They found that the variability of teacher effectiveness increases across grades and is most prominent in math. In an extreme case, they found a cohort of fifth grade students who had highly ineffective teachers in grades three through five. Those students scored about 50 percentile points below students with comparable previous achievement who had highly effective teachers for the same grades (Sanders & Rivers, 2001).

Rivkin et al. (2005) conducted a longitudinal study in the state of Texas. Researchers studied three cohorts of students as they moved from grade 3 to grade 7. The results on over 500,000 students in over 3,000 schools revealed that there was a positive relationship between teacher quality and student achievement (Rivkin et al., 2005).

Hanushek and Kimbo (2000) used three decades' worth of scores from six international tests of student achievement in math and science to measure the relationship between the quality of human capital and per capita growth. Thirty nine

countries provided direct observation of cognitive skills based on the battery of international tests. Thirty one countries provided measures of economic development. When looking specifically at school quality measures and their impact on per capita growth, Hanushek and Kimbo (2000) found that school spending had no impact on per capita growth. They also found that higher teacher quality impacted student learning as evidenced by higher test scores in math and science. According to their research, students with a thorough understanding of math and science go on to become engineers and scientists. The research also suggests that a good understanding of math and science determines individual productivity and income. Hanushek and Kimbo (2000) concluded that teacher quality had a “consistent, stable and strong relationship” with increased student achievement in science and math and increased earning power and productivity.

Clearly, teacher quality is important for student learning. Teacher quality impacts academic progress more than any other variable including socioeconomics, parental level of education, race, and individual school. Teacher quality has a positive impact on disadvantaged students and African American students and therefore has the potential to help close the historical achievement gap between poor, minority, urban students and middle class, white suburban students.

The effects of teacher quality go beyond one school year. The effects of teacher quality are cumulative and can last for up to four years. Nationally and internationally, teacher quality has impacted not only student learning, but earning potential and productivity as well.

Not only is there research to suggest that teacher quality impacts student achievement, there is also research to support the idea that first year teachers simply do not have the knowledge and skills needed to have a strong impact on student outcomes. There are some stories of first year teachers who appear to have a natural gift of teaching and do extremely well in the classroom. Unfortunately, for every one of those stories, there are many more that describe first year teachers who struggle through the school year with little to no focus on student achievement. It is not surprising that research indicates that first year teachers are dramatically less effective than more experienced teachers (Hanushek et al., 2005).

McArthur (1979) found that beginning teachers can be so anxious about their day- to-day survival that they lose conviction that they can make a difference in student learning. As a result, they develop a custodial attitude toward their students where their purpose is to provide protective supervision; to watch over and safeguard. Beginning teachers also may feel professionally “isolated” and focus only on survival (Griffin, 1987). Griffin (1987) found that when teachers learn through trial and error alone, their growth is limited. This “trial by fire” approach creates a passive teaching style in which teachers simply offer instruction to students with little focus on how the students receive the instruction (Feiman-Nemser, 1992).

It is understood that in order to become a strong teacher, one must endure the first year. There is simply no way to skip the difficult first year and jump straight into the more experienced years. The lessons learned at the beginning of a teaching career are invaluable and difficult to teach in a textbook. There is a way to provide support and training for beginning teachers so that they will remain in the profession and have

a stronger impact on student achievement in the future. One answer lies in providing support to teachers before they even begin their career.

There is also a way to provide a roadmap to high quality teaching. That roadmap is through teaching standards that define and guide high quality teaching. Other professions train their apprentices to meet very clear standards. You will not hear a medical resident say to a patient “I would like to try a new surgical technique that seems really fun. It hasn’t been fully researched yet, but I think that I can make it work.” You would not hear a new pilot announce to the passengers in the cabin “I would like to try a new landing today. I thought about it last night. I haven’t tried it before, and I haven’t talked it over with my colleagues, but it just might work.” You would not want a dentist to try a new root canal procedure with no data to back the effectiveness. You would never agree to an accountant using his/her own creative principles of accounting. You rest assured that the accountant will use the generally accepted principles of accounting; the widely accepted standards of the profession. Although these examples may seem silly, it is the way that many beginning teachers perform in the classroom. It is also the way that many preservice teachers begin their careers. They try new techniques that “feel” good or activities that students enjoy, without ever really knowing what are best practices. Hopefully, throughout their first few years, they stumble across some best practices, attend professional development, and learn from colleagues. Hopefully, they grow from these haphazard, often spontaneous experiences.

All professions have standards that guide their work. This applies to the field of education as well. There are clear, widely accepted teaching standards that provide

a roadmap to preservice and beginning teachers. It is important to share that roadmap. Teacher preparation programs are extremely important in this process. They have the opportunity to provide the best training possible and to help preservice teachers to identify characteristics of teacher quality and standards for the profession. Teacher preparation programs not only help teachers identify characteristics of teacher quality but also provide actual measures that will help preservice teachers learn and grow.

One critical component of the preservice teacher phase is evaluation. It is important to provide preservice teachers with an explanation of the evaluation criteria that will be used based on accepted teaching standards. It is also critical to observe preservice teachers in the classroom setting to provide measures of those characteristics as well as feedback on strengths and weaknesses. Observations should take place throughout the student teaching experience and provide opportunities for dialogue before and after each observation. This critical component of the preservice program provides an opportunity to assess teaching standards and promote professional development.

Statement of the Problem

In order to improve student learning, you must improve teacher quality. In order to improve teacher quality, you must know what teacher quality looks like. Identifying characteristics of teacher quality and providing measurable feedback on those characteristics can improve the knowledge, skills and dispositions of preservice teachers and help them to be more effective in the classroom. A problem in the field of education is that without the use of standards, it can be very difficult to measure teacher quality. Without standards, trying to measure teacher quality is much like

Justice Potter Stewart trying to define obscenity. "I shall not today attempt further to define the kinds of material I understand to be embraced . . . [b]ut I know it when I see it." (Jacobellis v. Ohio, 1964).

It is not enough for educators to be able to identify quality teaching when they see it. They must be able to share those characteristics with beginning teachers. The characteristics must be measurable. The measures must be discussed in a way that promotes professional development. Identification of characteristics, measurement of characteristics and discussion of characteristics in a way that promotes professional development are essential to any evaluation program. There is a need for a valid and reliable evaluation instrument that is built on teaching standards.

Purpose of the Study

The Student Teaching Assessment Rubric (STAR) is built on widely accepted teaching standards and used with preservice teachers. It was designed to evaluate preservice teachers' knowledge, skills and dispositions in the classroom. The purpose of this study was to examine the underlying structure of the STAR.

Research Questions

Two research questions were examined in this study.

Question 1

Does the specified model based on INTASC standards fit the empirical data from the STAR?

Question 2

Are the parameter estimates statistically significant from zero?

Delimitations of the Study

There were delimitations in this study. First, this study only examined preservice teachers, not teachers in the field. Only the factors associated with the INTASC standards were evaluated. Furthermore, this study was limited to STAR data at one university. These delimitations may influence the generalization of the results.

Limitations of the Study

There were limitations to the study that were inherent to the methodology. Data used in this study came from three cohorts of university students. These limitations may influence the generalization of the results.

Assumptions and Operational Definitions

The study included a number of assumptions about preservice teacher performance evaluation and operational definitions of terms used in the literature on teacher quality.

Assumptions

One primary assumption was that quality teaching is directly related to student achievement. The most efficient way to improve learning in the classroom is to improve teaching. Therefore, the primary focus of school reform should be teacher quality.

The second assumption was that teacher evaluation can improve teaching and help preservice teachers transition into quality teachers. The teaching style of novice teachers can be greatly shaped by evaluation techniques. Preservice teachers can practice and hone their skills in the classroom based on the feedback of evaluations

from experienced teachers and university supervisors. Preservice teachers can develop professional dialogue and problem solving strategies through evaluation conversations.

The last assumption was that it is very difficult to measure teacher quality. Teaching is a complex profession that requires sophisticated knowledge and skills in many areas. Individuals use a variety of variables to define teacher quality including teacher test scores, content knowledge, pedagogical knowledge, advanced degrees, academic major, teaching experience, and traditional certification versus alternative certification. These variables provide a weak link to student achievement and little insight for preservice teachers. There is a need for a preservice teacher evaluation instrument that is reliable and valid.

Operational Definitions

- **Preservice teacher** was the term used to describe university students in the final phase of their teacher preparation program, the student teaching experience. Preservice teachers worked in a classroom under the supervision of a cooperating teacher (usually a veteran teacher with years of experience) and a university supervisor (usually a professor who has worked closely with the student in the teacher preparation program).
- **Teacher quality** can be difficult to define. Some simply refer to teacher quality as successful teaching based on outcomes. Others argue that there is much more to quality teaching than simple student outcomes. Preservice teachers especially need more guidance in the definition of teacher quality. This research was guided by a definition that uses evidence based on observations and related to teaching standards.

Quality teaching ... is about more than whether something is taught. It is also about how it is taught. Not only must the content be appropriate, proper, and aimed at some worthy purpose, the methods employed have to be morally defensible and grounded in shared conceptions of reasonableness. (Fenstermacher & Richardson, 2005; p. 189).

The “shared conceptions of reasonableness” lie in teaching standards. For the purpose of this study, teacher quality was defined as what teachers should know and be able to do, according to the standards set by the Interstate New Teacher Assessment and Support Consortium (INTASC).

- **Student Teaching Assessment Rubric (STAR)** is a direct observation tool based on INTASC standards that is used to evaluate preservice teachers during their student teaching phase.
- **Interstate New Teacher Assessment and Support Consortium (INTASC)** is a consortium of state and national education agencies working collaboratively to focus on the needs of beginning teachers, including providing standards to guide beginning teachers.
- **INTASC Standards** are standards that are performance based describing what beginning teachers should know and be able to do.

Significance of the Study

Research shows that variables related to teacher quality are more strongly related to student achievement than student demographics, class size, overall spending levels, and teacher salaries (Darling-Hammond, 2000; Sanders & Rivers, 2001).

Knowing that teacher quality is directly related to student achievement, it is crucial to have an effective way to measure teacher quality. This is especially important when measuring the quality of preservice teachers as they develop.

Performance evaluation programs for preservice teachers can deeply impact the profession in several ways. Preservice teacher evaluation programs can directly impact teaching styles and orientations (Veenman, 1984). Preservice teacher evaluation programs can also positively impact teachers' commitment to and desire to stay in the profession (Rosenholtz, 1986). Another benefit of preservice teacher evaluation is the formal or informal assistance that preservice teachers receive as a result of the preservice teacher evaluation program (Sclan & Darling-Hammond, 1992).

Preservice teachers face many frustrations. Despite some of the longest days of teaching, there is a light at the end of the tunnel. Many teachers are able to navigate their way through the preservice stages and go on to be very effective teachers. Unfortunately, many preservice teachers become overwhelmed and walk away. A strong evaluation tool that can be used at the very beginning of a teacher's career can greatly impact the quality of the teacher as well as the profession as a whole. There is a need for a valid and reliable evaluation tool. A study of the STAR will add valuable insight to this field of study.

Design and Overview of the Study

This study was designed to examine the underlying structure of the STAR. An archival data set of responses from university supervisors of student teachers was used. Confirmatory factor analysis was conducted to examine the research questions.

Chapter II includes a review of research related to measures of teacher quality including subject matter knowledge, advanced degrees, pedagogical knowledge, teacher test scores, and level of certification. There will also be a section on various methods of preservice teacher evaluation and widely acceptable teaching standards,

including INTASC standards. Chapter III addresses the methodology of the study. Chapter IV discusses the key findings, and Chapter V discusses the implications and suggestions for future research.

Summary

School districts across the country are clear about the mandate they are facing: fill the classrooms with teachers who are highly qualified. School districts are also clear about the challenge: do this in spite of the largest teacher shortage in the history of the profession. School districts understand that the single most important factor in student achievement is teacher quality. Despite this wealth of knowledge about the situation at hand, school districts continue to struggle to meet the mandate. While there is a great deal of data related to the importance of teacher quality, there continue to be unanswered questions.

Why and how do some beginning teachers foster high student achievement while others fall short? Are there certain characteristics of teacher quality that can be measured? If so, what are the characteristics of teacher quality, and how can the profession groom preservice teachers in that direction? How can the profession assess teacher quality and provide support and development needed to positively impact student learning?

Knowing that teacher quality is the single most important factor in student achievement is only half of the equation. The other half is finding a way to identify and measure teacher quality. School districts and teacher development programs across the country cannot begin to improve the quality of teaching and improve

student learning until there is a way to assess teacher quality and to provide support and professional development to grow quality teachers.

The growth of the profession begins with the preservice teacher. The preservice teaching experience provides fertile soil for growth and development. The evaluation arm of a preservice teacher preparation program is equally important in fostering and developing quality teachers.

Educational research has revealed a major need for studies that contribute to the knowledge base on methods of assessing quality in preservice teachers (Borko, Liston & Whitcomb, 2007; Cochran-Smith, 2006; Darling-Hammond, 2006, Loughran, 2007; and Zeichner, 2007). This study is an attempt to add to the body of knowledge.

CHAPTER II: REVIEW OF LITERATURE

The issue of school reform has a long, rich history. As far back as 1840, Horace Mann led the “common school” movement as a way to reform education and improve society. In more recent history, the National Commission on Teaching and America’s Future published *What Matters Most: Teaching for America’s Future* in 1996. This publication focused on teacher quality as the subject of school reform. The authors argued that the single most important factor for achieving America’s educational goals was to focus on the teacher – including recruitment, preparation, and support. The commission’s goal was to make sure that every classroom across the country had a caring, competent and qualified teacher who had the knowledge and skills to teach all students to learn. Equally important was the focus on developing school system programs that could support teachers in this work. The commission had three premises:

1. What teachers know and can do is the most important influence on what students learn.
2. Recruiting, preparing and retaining good teachers is the central strategy for improving our schools.
3. School reform cannot succeed unless it focuses on creating the conditions in which teachers can teach, and teach well (National Commission on Teaching America’s Future, 1996).

The commission also set the following goal: “By the year 2006, we will provide every student in America with what should be his or her educational birthright: access to competent, caring, qualified teachers in schools organized for success (National Commission on Teaching America’s Future, 1996).

This report was very influential in creating a focus on teacher quality as a factor in student achievement. It provided a research-based argument, including 126 references, which linked teacher quality to student achievement. According to the report, “research shows that teacher knowledge of subject matter, student learning and teaching methods are all important elements of teacher effectiveness” (Darling-Hammond, 2000).

While more recent studies have also found that the single most important school-level factor associated with student learning is the quality of the classroom teacher (Aaronson et al., 2007, Rockoff, 2004, Rivkin et al., 2005, Kane, Rockoff & Staiger, 2006), there continues to be uncertainty regarding how to appropriately measure teacher quality.

Physicians are measured by national standards set by the American Medical Association. Lawyers are measured using national standards from the American Bar Association. Generally Accepted Accounting Principles (GAAP) serve as the framework of standards for financial accounting.

Dictionaries give two uses of the term *standard*; to rally and to measure. Both are appropriate when talking about teaching standards. Teaching standards are the educational principles and values that the profession is built upon. Teaching standards provide a vision of quality teaching and quality learning that guide the profession.

Teaching standards are also used to measure. Educators use teaching standards to make judgments about professional performance (Fenstermacher & Richardson, 2005).

There are four sets of standards that guide the teaching profession, created by these professional organizations: the National Board for Professional Teaching Standards, the National Council for Accreditation of Teacher Education, the American Association of Colleges for Teacher Education, and the Interstate New Teacher Assessment and Support Consortium. Below is a brief overview of each set of standards. There is a more in-depth overview of the INTASC standards because they are the standards developed for preservice teachers.

National Board for Professional Teaching Standards

The National Board for Professional Teaching Standards (NBPTS) was formed in 1987 to advance the quality of teaching and learning in the United States by developing professional standards for accomplished teaching. The National Board also created a voluntary system to certify teachers who meet those standards and to then integrate certified teachers into educational reform efforts. The organization is an independent, non-profit, non-partisan and non-governmental national organization with a broad membership base that includes teachers, state governors, school administrators, teacher unions, school board leaders, college and university officials, business executives, foundations and concerned citizens. The NBPT standards are the most highly regarded standards for measuring highly accomplished teaching.

The standards, developed by teachers from all grade levels and subject areas through their professional associations, aim to capture substantive knowledge about

teaching and learning. The standards are performance based; describing exactly what teachers should know and be able to do. These standards are generally applied to experienced teachers rather than beginning teachers because of the advanced experience, knowledge and skill required to meet them.

There are five core propositions that provide the foundation for the knowledge, skills, dispositions and beliefs that characterize National Board Certified Teachers.

The core propositions are:

1. Teachers are committed to students and learning.
2. Teachers know the subjects they teach and know how to teach those subjects to students.
3. Teachers are responsible for managing and monitoring student learning.
4. Teachers think systematically about their practice and learn from the experience.
5. Teachers are members of learning communities.

In addition to the general standards that apply to all teachers, there are specific standards for 25 different grade levels and subject areas. To gain National Board Certification, teachers must demonstrate mastery of the standards appropriate to their teaching area.

National Council for Accreditation of Teacher Education

The National Council for Accreditation of Teacher Education (NCATE) is a non-profit, non-governmental alliance of 33 national professional education and public organizations that has set programmatic standards for colleges and universities that offer academic programs to prepare teachers and other school professionals. The

number of NCATE accredited institutions has risen from 492 in 1999 to 632 in 2007. Currently 78 colleges of education are seeking NCATE accreditation. NCATE accreditation is a mark of distinction and provides recognition that the college of education has met national professional standards for the preparation of teachers and other educators.

NCATE was founded in 1954. Five groups were instrumental in the creation of NCATE: the American Association of Colleges for Teacher Education (AACTE), the National Association of State Directors of Teacher Education and Certification (NASDTEC), the National Education Association (NEA), the Council of Chief State School Officers (CCSSO), and the National School Boards Association (NSBA). NCATE actually replaced AACTE as the agency responsible for accreditation in teacher education. Recognizing the need for a strong, independent, quality assurance mechanism composed of all key stakeholders in education, these five groups came together to promote quality teaching. The mission of NCATE is

to help establish high quality teacher, specialist, and administrator preparation. Through the process of professional accreditation of schools, colleges and departments of education, NCATE works to make a difference in the quality of teaching, teachers, school specialists and administrators. NCATE believes every student deserves a caring, competent, and highly qualified teacher (p. 1).

NCATE accredited schools produce over 2/3 of the nation's new teacher graduates (data compiled by Westat, computed by NCATE). Graduates of NCATE accredited colleges of education pass the Educational Testing Service (ETS) subject matter and pedagogy examinations at a higher rate than graduates of unaccredited colleges of education (Educational Testing Service, 1999). Approximately 70% of the

189 doctoral granting institutions are NCATE accredited or candidates for accreditation.

NCATE has four major goals. The first goal speaks directly to teaching standards. NCATE's first goal is to maintain high standards for the knowledge, skills and professional dispositions required of educators and for the units and programs that prepare them to practice. This goal supports NCATE's belief that a necessary element that qualifies an occupation as a profession is the consensus regarding the standards of the profession. The NCATE standards are divided into three sections: unit standards, program standards and standards for professional development schools.

NCATE also has unit standards. Unit refers to the college of education. These are standards that guide the college's direction for programs, courses, teaching, candidate performance, scholarship, service and accountability. Unit standards provide the bases for the college of education's intellectual philosophy and institutional standards.

Each college of education seeking accreditation from NCATE is required to submit a conceptual framework in order to be reviewed. The conceptual framework provides the shared vision for the institution. Faculty members are expected to develop the conceptual framework in collaboration with the members of the professional community. The conceptual framework should reflect the institution's commitment to diversity, integration of technology, and maintaining state standards. A team of Board of Examiners looks for evidence of the conceptual framework during their visit. The team uses a rubric to evaluate the institution based on the unit standards.

NCATE's Specialty Areas Studies Board has approved national standards for 20 program areas. These guidelines were developed by professional associations that are members of NCATE. NCATE program standards are revised every seven years.

The program areas are listed below:

National Association for the Education of Young Children
 Association for Educational Communications and Technology
 International Society for Technology in Education
 Educational Leadership Constituent Council
 Association for Childhood Education International
 Teachers of English to Speakers of Other Languages
 National Council of Teachers of English
 North American Association for Environmental Education
 American Council on the Teaching of Foreign Languages
 National Association of Gifted Children/Council for Exceptional Children
 American Alliance for Health, Physical Education, Recreation and Dance/American Association for Health Education
 National Council of Teachers of Mathematics
 National Middle Schools Association
 International Reading Association
 American Library Association/American Association of School Librarians
 National Association of School Psychologists
 National Science Teachers Association
 National Council for Social Studies
 Council for Exceptional Children
 International Technology Education Association/Council on Technology Teacher Education

American Association of Colleges for Teacher Education

The American Association of Colleges for Teacher Education (AACTE) is a national alliance of educator preparation programs committed to the highest quality of professional development of teachers and school leaders to support P – 12 student learning. AACTE membership includes over 800 organizations representing public and private colleges and universities in every state, the District of Columbia, the

Virgin Islands, Puerto Rico, and Guam. AACTE strives to improve professional development and student learning through advocacy, leadership and service.

AACTE's Standards Based Teacher Education Project (STEP) focuses on institutional accountability for the preparation of teachers who can help P – 12 students meet academic goals. STEP focuses on the redesign of teacher preparation programs by aligning with national and state academic content standards and professional teaching standards. STEP is nationally recognized and has worked with teachers and teacher educators in over 45 programs in seven states for the past eight years. STEP is based on three principles:

- 1) Teachers must know the subjects they are teaching.
- 2) Teachers must know how to teach students to learn at high levels.
- 3) Teachers must know how to monitor and assess how well students are learning.

Interstate New Teacher Assessment and Support Consortium Standards

The Interstate New Teacher Assessment and Support Consortium (INTASC) was established in 1987 to provide an opportunity for states to focus on the needs of beginning teachers and to collaboratively rethink teacher assessment, licensure and induction. Members of the consortium include state education agencies and national educational organizations committed to the reform of teacher preparation, licensing and on-going professional development. INTASC established a task force to consider the kind of changes that needed to be made. The group wanted to create standards that embody the kinds of knowledge, skills, and dispositions that beginning teachers need

to practice responsibly when they enter teaching and that prepare them for eventual success as Board-certified teachers later in their careers.

The INTASC task force decided to begin its work by articulating standards for a common core of teaching knowledge and skills that should be acquired by all preservice teachers. It was important to start with this core knowledge in order to develop a common commitment to ethical practice and knowledge that provides the glue that holds members of a profession together, creates a common language and develops a set of understandings and beliefs that permit professional conversation similar to other professions.

The standards are performance based in that they describe what teachers should know and be able to do. The standards were based on the National Board five major propositions:

1. Teachers are committed to students and their learning.
2. Teachers know the subjects they teach and how to teach those subjects to diverse learners.
3. Teachers are responsible for managing and monitoring student learning.
4. Teachers think systematically about their practice and learn from experience.
5. Teachers are members of learning communities.

The expectation is that preservice teachers have at least an awareness of the kinds of knowledge and understanding needed, as well as the resources available, to develop these skills. Preservice teachers should also have some capacity to address the many facets of curriculum, classroom, and student life, and must have the dispositions

and commitment that pledge them to professional development and responsibility. These standards aim to develop preservice professionals while contributing to the development of the profession.

An in-depth summary of the standards can be found in Table 1. More information about INTASC and the standards can be found on the Council of Chief State School Officers (CCSSO) website.

Table 1

INTASC Teaching Standards

<u>Standard</u>	<u>Description</u>
Content Pedagogy	The teacher understands the subject matter and is able to communicate the curriculum to the students in a way that is relevant, meaningful and understandable. The teacher understands that the subject is not fixed, but ever changing. The teacher understands his or her responsibility to stay abreast of changes in the subject. The teacher uses multiple strategies, viewpoints and a variety of resources to convey the subject to the students. The teacher encourages students to question the subject and to test hypotheses.
Student Development	The teacher understands child development in all 5 domains (physical, social, emotional, moral and cognitive). The teacher appreciates variations in student learning and adapts to meet individual student needs. The teacher celebrates students' strengths and sees their weaknesses as an opportunity for growth. The teacher meets students where they are and builds on their strengths and experiences.

Diverse Learners

The teacher understands diversity related to learning styles, multiple intelligences, and performance modes. The teacher understands the needs of students with disabilities, English as a second language students as well as students from diverse cultural backgrounds. The teacher believes that all students can learn and maintains high expectations for all students. The teacher shows respect for all students and makes all students feel valued for their strengths and differences. The teacher uses multiple instructional strategies to meet the needs of all students. The teacher is aware of community resources available to students and their families.

Instructional Strategies

The teacher uses a variety of instructional strategies that promote critical thinking, problem solving and active inquiry. The teacher uses a variety of instructional strategies, technology and work products. The teacher constantly monitors and makes adjustments based on observation.

Motivation and Management

The teacher uses general knowledge of psychology, anthropology and sociology to develop a learning environment that is conducive to learning. The teacher understands the complex nature of classroom management and uses a variety of strategies to promote positive relationships and cooperation in the classroom. The teacher manages time and materials well. The teacher is committed to democratic values in the classroom.

Communication/Technology

The teacher understands communication theory, the role of language in learning and the power of language in self expression and identity development. The teacher uses a variety of communication tools including audiovisuals and computers to enrich learning.

Planning	The teacher is able to develop long-term and short-term plans in order to meet curriculum goals as well as the diverse needs of the students. The teacher is able to make adjustments to the plan based on student response. The teacher understands that planning is a collegial activity.
Assessment	The teacher uses different types of assessments including criterion-references and norm-referenced instruments, traditional standardized and performance-based tests, observation systems and assessments of student work. The teacher understands basic measurement terms such as validity, reliability, bias, and scoring concerns. The teacher uses ongoing formal and informal assessment to monitor and identify student strengths. The teacher allows students to use self assessment instruments. The teacher maintains assessment records.
Reflection and Professional Development	The teacher reflects on the learning environment and makes adjustments based on the reflection. The teacher is willing to seek and give help. The teacher actively seeks out professional development opportunities.
School/Family/Community	The teacher understands organizational development as related to the school. The teacher is able to work well in the classroom as well as in the larger school community. The teacher also understands how the community can impact the school and the classroom. The teacher understands general school law as well as ethical teaching standards. The teachers understands the “whole child” and is willing to work with parents, school staff and community staff to meet the needs of the student while maintaining respect and confidentiality of the student. The teacher

contributes to the goals of the school and community.

All of the standards mentioned above focus on what teachers should know and be able to do. The standards address the knowledge, skills, and dispositions needed to be successful in the classroom, regardless of the subject area. Local, state and national educational leaders have come together to agree on how teachers should be prepared, what they should know at the preservice stage, as well as how they should develop into experienced teachers.

It would make sense to measure educational professionals using standards as other professions have done, and this study will examine one instrument that attempts to do so. However, when attempting to measure quality teaching, other researchers have looked to demographic variables such as general academic ability and intelligence, subject matter knowledge, knowledge of teaching and learning, advanced degree, teaching experience and type of certification. A review of this research is below.

Measures of General Academic Ability

It may be logical to assume that a teacher with strong general academic ability, intelligence, pedagogical knowledge and subject matter knowledge would perform better than a teacher of average academic ability and intelligence. This assumption goes as far back as 1940 when researchers attempted to show a correlation between teacher performance and teacher intelligence. Studies did yield a positive correlation between teacher performance and teacher intelligence (Hellfritsch, 1945; LaDuke, 1945; Rostker, 1945; Skinner, 1947); however the relationships were small and statistically insignificant. Later studies showed that there is no relationship between

measured teacher intelligence and student achievement (Schallock, 1979; Soar, Medley, & Coker, 1983).

There are also several studies that examine various teacher test scores including state licensing exams such as the Praxis, American College Testing (ACT) exam, and Tennessee Value Added Assessment System (TVAAS). A review of that research along with a description of the tests is below.

Most states require that preservice teachers pass a licensing exam before they can become a licensed teacher. It is natural to assume that preservice teachers who score high on licensing exams would be more knowledgeable than a teacher with lower test scores and thus more effective in the classroom. The research on teacher licensing test scores is mixed.

Ronald Ferguson (1991) analyzed test scores in 900 school districts in the state of Texas. He looked at teacher quality as measured by scores on the state licensing examination, a Master's degree, and years of teaching experience. He found that 40% of the variance in student test scores could be attributed to teacher scores on the licensing examination.

Ferguson and Ladd (1996) explored the relationship between teachers' American College Testing (ACT) exam scores and 3rd grade students' gains. The ACT exam assesses high school students' general educational development and their ability to complete college level work. The ACT scores combine English, mathematics, social studies, reading and natural sciences. Ferguson and Ladd (1996) found that 3rd to 4th grade reading gains were positively related to the teacher's ACT score. The relationship was not clear for math scores.

While examining administrative data from North Carolina to explore the relationship between student achievement and advanced degrees, Clotfelter, Ladd, and Vigdor (2006) also examined the impact of teacher licensure test scores. Clotfelter et al. (2006) concluded that there was a positive correlation between teacher licensure test scores and student achievement. These effects were large for math but much smaller for reading.

Using a unique data set of elementary students in North Carolina to link teachers to their individual students, Goldhaber (2007) used a value-added gain score model with lagged test scores as a regressor to examine the effects of teacher licensure test scores on student achievement. Goldhaber (2007) used scores from the ETS Praxis I test. Goldhaber (2007) found that there was a small positive relationship between teacher licensure test scores and student achievement.

Buddin and Zamarro (2008) used a value added approach to examine the student achievement of students in grades 2 through 5 from the Los Angeles Unified School District for five consecutive years (2000 to 2004). The students were enrolled in self-contained classrooms taught by a single teacher. California requires new elementary teachers to pass up to three tests to receive state certification (basic skills, subject matter knowledge and reading pedagogy). Buddin and Zamarro (2008) used teacher license test scores on all three tests from six cohorts of teachers in the California State University system. Based on their data, Buddin and Zamarro (2008) concluded that teacher licensure scores on the tests mentioned above have little if any effect on classroom student achievement.

State licensing exams, the American College Testing exam and the Tennessee Value Added Assessment System all serve to assess what a teacher knows.

Demographic variables such as teacher's test performance do not correlate positively with effective teaching or high student achievement. Professional standards provide a better measure of what teachers know. INTASC standards identify exactly what beginning teachers should know and be able to do in the classroom. These standards would be much more beneficial in identifying teacher quality than individual tests of academic ability or intelligence.

Many consider a teacher's academic major to represent the depth of their subject matter knowledge. One would imagine that a teacher who majored in the subject he or she is teaching would be more effective than a teacher who did not. Since elementary school teachers are responsible for teaching all content areas, this variable has greater importance for middle and high schools. As expected, a teacher's academic major has proven to have little to no impact in elementary school. However, in middle school and high school, a teacher's academic major in math and science has a positive impact on student achievement. There is little to no impact in the area of language arts/English and social studies/history (Wenglinsky, 2000, Frome, Lasater, & Cooney, 2005).

There have been several studies that have attempted to establish a relationship between teacher's scores on the National Teacher Examination (NTE) and teacher performance and student achievement. The NTE is designed to assess subject matter knowledge and pedagogical knowledge and is used for teacher licensure and certification in many states. The results of these studies find that there is no consistent

relationship between teacher NTE scores and student outcomes. Small statistically insignificant relationships, both positive and negative were found (Andrews, Blackmon & Mackey, 1980; Ayers & Qualls, 1979; Haney, Madaus, & Kreitzer, 1986; Quirk, Witten, & Weinberg, 1973).

One example of a study on NTE scores as a measure of teacher quality is Summers and Wolfe's (1975, 1977) study in Philadelphia. Summers and Wolfe (1975, 1977) conducted a study of 627 sixth graders in Philadelphia. They attempted to examine if students learn more from teachers who scored high on the NTE Common Examination, a test used specifically to measure subject matter knowledge. They found that students actually learned less from teachers who scored high on the NTE Common Examinations. They explained their surprising findings by pointing out that student achievement is the result of a complex mix of socioeconomic status, teacher quality, school quality and peer group characteristics. They concluded that appropriate measures of teacher quality have yet to be discovered.

Byrne (1983) analyzed 30 studies that related teacher subject matter knowledge to student achievement. Byrne's (1983) measures of subject matter knowledge were either standardized subject matter knowledge tests or number of college courses taken in a subject area. Of the 30 studies, 17 showed a positive relationship, while 14 showed no relationship. Ashton and Crocker (1987) conducted a similar analysis and determined that very few studies showed a positive relationship between teacher subject knowledge and student achievement.

As mentioned above, subject matter knowledge is of greater important for middle school and high school teachers who focus on one specific subject. It is also of

greater importance since subject matter in middle and high schools becomes more complex and sophisticated. Subject matter knowledge is particularly important in the area of science and mathematics.

Rowan, Chiang and Miller (1997) used data from the National Educational Longitudinal Study of 1988 (NELS: 88). The data included student test scores in math from the spring of their 8th grade year and the spring of their 10th grade year. Tenth grade teachers took a survey including a single high school mathematics test item. The researchers controlled for whether or not the teacher had a mathematics-related degree. Rowan, Chiang, and Miller (1997) found that the students of teachers who answered the math question correctly experienced larger gains from 8th grade to 10th grade. These findings would carry much more weight if the teachers had answered more than one mathematics question.

A much more rigorous study was conducted by Monk and King (1994) in the area of math and science. Monk and King (1994) used data from the Longitudinal Study of American Youth. Using data from 2,829 students, Monk and King (1994) found that teachers' subject matter knowledge as measured by coursework in the subject field had a positive relationship to student achievement in math and science.

Using the same data set, Monk and King (1994) found a positive and negative relationship between teacher subject matter knowledge and student achievement. All relationships were statistically insignificant. There was evidence of cumulative effects of teacher subject matter knowledge and student achievement in math. If a student had a math teacher who was poorly prepared in the subject area, the negative impact on

their student achievement lasted for several years. The results differed for high achieving students and low achieving students and for different grades.

In review of 65 science teachers, Druva and Anderson (1983) found that the students' science achievement was positively related to teachers' subject knowledge as measured by courses taken. Hawk, Coble and Swanson (1985) found similar findings in mathematics. Both Druva and Anderson (1983) and Hawk, Coble and Swanson (1985) found that the difference in student achievement was higher in high levels of math and science.

Few would argue against the notion that it is important for teachers to have a good understanding of the subject he or she teaches. However, using courses taken or subject matter standardized test scores does not appear to be an appropriate or useful measure of teacher effectiveness. It would be more appropriate to use a standards-based instrument to measure teacher effectiveness and its impact on student achievement.

INTASC standard 1 specifically addresses subject matter knowledge. INTASC standard 1 addresses content area: "The teacher understands the central concepts, tools of inquiry, and structures of the discipline(s) he or she teaches and can create learning experiences that make these aspects of subject matter meaningful for students.". INTASC also provides descriptors to measure this area. The standards include the teacher's ability to understand the subject matter and to communicate it to students in a way that is relevant, meaningful and understandable. The standards also include the teacher's ability to see the subject matter as ever-changing and to honor his or her responsibility to remain current and to stay abreast of the changes. INTASC

recognizes that the teacher has a responsibility to use multiple strategies, viewpoints and resources to encourage students to question the subject and to test their hypotheses. The INTASC standards would be a much more appropriate measure of teacher effectiveness.

Knowledge of Teaching and Learning

Researchers recognize that it is not only important for teachers to understand the subject matter; they must also understand how to teach the subject and how students learn. Some have tried to establish a relationship between a strong understanding of teaching and learning and student achievement.

Ashton and Crocker (1987) reviewed studies of teacher quality and found that four of the seven studies found a positive relationship between education coursework and teacher performance. This was larger than the number that showed a relationship between subject matter knowledge and teacher performance.

Evertson, Hawley and Zlotnik (1985) reviewed research of teacher preparation programs in an attempt to make recommendations for improvement. The researchers paid special attention to research on teacher preparation and teacher effectiveness. They found a positive relationship between teachers' formal education training and student achievement in 11 of 13 studies.

As with subject matter knowledge, teacher preparation in the area of how to teach math and science seems to be more significant. There have been several studies that have examined the relationship between teacher coursework in math and science methodology (courses that focus on how to teach math and science as opposed to courses that focus on the actual theories of math and science) and student

achievement. For example, Perkes (1967) found that there was a positive and statistically significant relationship between teachers' coursework in how to teach science and student achievement but there was no statistically significant relationship between teachers' coursework in pure science and student achievement. This supports the notion that knowing the fundamentals of the subject area is very different from knowing the methodology of teaching the subject area.

Using findings from the National Study of Mathematical Abilities, Beagle (1979) studied teacher education and student achievement. Beagle (1979) found that there was a strong relationship between teachers' number of mathematics methods course credits and student performance. Monk and King (1994) also found that teacher education coursework has a positive relationship with student achievement and that the relationship was stronger than the relationship between subject matter knowledge and student achievement.

Using data from more than 200 graduates of a single teacher education program, Ferguson and Womack (1993) studied the influences on 13 dimensions of teaching including education coursework, subject matter coursework, NTE subject matter test scores and GPA in the student's major. They found that education coursework accounted for more than four times the variance in teacher performance.

Guyton and Farokhi (1987) conducted a similar study using a standardized observation instrument to measure 12 dimensions of teacher performance for more than 270 teachers. They found a positive relationship between teacher education coursework and teacher performance in the classroom. The relationship between teacher performance and subject matter test scores was positive but insignificant.

It is important for teachers to understand the subject they teach. It is equally important for teachers to understand teaching methodology and to be able to communicate the subject to students in a way that is understandable and meaningful. A standards based approach is a much better way to measure understanding of teaching methodology. As research shows above, a simple survey of methodological courses taken is not enough to judge what a teacher knows and is able to do. It is also not enough to predict teacher performance in the classroom.

An example of a standards based approach to assessing teacher effectiveness is INTASC standards. INTASC Standards pay special attention to student learning (Standard 2), diverse learners (Standard 3), instructional strategies (Standard 4), learning environment (Standard 5), planning instruction (Standard 7), assessment (Standard 8), and reflection and professional development (Standard 9). INTASC Standards can be used as a guide to measure teacher effectiveness. The language and descriptors of the standards can reflect a teacher's understanding of child development and variations of instructional strategies needed to meet individual student needs. They also measure the teacher's understanding of diversity as related to learning styles, multiple intelligences, disabilities, English language deficiencies, and diverse cultural backgrounds. There are INTASC standards that measure the teacher's use of instructional strategies that promote critical thinking, problem solving and active inquiry. There are also standards that measure the teacher's general knowledge of psychology, anthropology, and sociology as it relates to developing a classroom environment that is conducive to learning. INTASC standards measure the teacher's ability to develop long term and short term lesson plans that meet the curriculum goals

as well as the needs of the students. There are also standards that measure the teacher's use of national, state and local assessment data as well as a general understanding of types of assessment including criterion-referenced and norm referenced instruments, traditional standardized and performance based tests, observation systems and informal assessment of student work. Overall, INTASC standards reflect the wide range of knowledge, skills and dispositions that are essential for an effective beginning teacher.

Use of INTASC standards to measure teacher effectiveness would provide much richer data than number of education courses taken. Use of the INTASC standards would also provide much greater feedback to teachers and support the teacher's professional development.

Advanced Degree

One may think that the attainment of a Master's degree would add to a teacher's body of knowledge and thus make him or her more effective. Research does not necessarily support this premise. Clotfelter et al. (2007) used 10 years worth of administrative data from North Carolina to explore the relationship between teacher characteristics and credentials and student achievement in math and reading. One teacher characteristic they explored was advanced degrees. They found that not only were advanced degrees not positively associated with student achievement, in some cases advanced degrees were negatively correlated with student achievement.

Betts, Zau and Rice (2003) used longitudinal data from San Diego to study impacts on student achievement. One part of that study was to examine factors related to student achievement and to identify factors that are not related to student

achievement. Betts, Zau and Rice (2003) found that having an advanced degree does not have a statistically significant impact on student achievement.

Aaronson et al. (2007) also found substantial variation when looking at teacher quality in the Chicago public schools. They used administrative data to determine the importance of teacher characteristics of quality on student achievement in math. One teacher characteristic they explored was having an advanced degree. Having an advanced degree did not have a positive statistically significant impact on math student achievement. In some cases, the impact was negative. Aaronson et al. (2007) found that traditional demographic measures explain little of the total variation in teacher quality.

Hanushek et al. (2005) used a semi-parametric approach to measure quality on the basis of value added to student achievement in a large urban school district. They examined observable teacher characteristics including race, experience and educational preparedness. Understanding that a student's academic achievement is impacted by not only educational inputs but also by past history of family, neighborhood, and previous school experience, Hanushek et al. (2005) decided to focus on student achievement gains rather than levels. They found substantial variations in teacher quality. Teacher characteristics that may seem obvious in impacting student achievement, such as having an advanced degree, did not have a statistically significant positive impact on student achievement gains. In some cases, there was a negative impact. Furthermore, when looking at characteristics of teacher quality in general, they found substantial variations.

Rivkin et al. (2005) used data from the UTD Texas Schools Project to examine whether or not student achievement in reading and math was impacted by observable teacher characteristics such as educational level. Rivkin et al. (2005) used data from three cohorts of 3rd through 7th grade student test scores from over 200,000 students in over 3,000 public elementary and middle schools. This large data sample permitted more precise estimates of average test scores and test score gains. Rivkin et al. (2005) found that having a Master's degree did not raise the quality of teaching. All effects were statistically insignificant.

Buddin and Zamarro (2008) used a value added approach to examine teacher characteristics and student achievement. They examined the relationship between student achievement and having an advanced degree. They found that student achievement is unaffected by whether the classroom teacher has an advanced degree.

Jacob and Lefgren (2008) used data from a midsized school district to examine the relationship between teacher characteristics and student achievement. One of the characteristics they examined was advanced degrees. They found that having an advanced degree has only a small effect on student achievement.

Teaching Experience

It seems obvious that a teacher's years of experience would have a positive impact on student achievement. As teachers grow in the profession, one would expect them to develop more strategies to meet the needs of students. However, research on the impact of teaching experience on student achievement is mixed.

Nye et al. (2004) used a hierarchical linear model to examine Stanford Achievement Test scores in reading and math from Kindergarten through 3rd graders.

They examined the relationship between teaching experience and student achievement. They found that the estimated relationship between teacher experience and student achievement gains was substantial but only statistically significant for 2nd grade reading and 3rd grade math.

Rivkin et al. (2005) used UTD Texas Schools Project data to examine not only the relationship of advanced degrees on student achievement but teaching experience as well. They found that it was difficult to measure teacher quality using observable characteristics. The data revealed that it was difficult to measure the impact of teaching experience on student achievement.

When studying data from elementary students and teachers in North Carolina, Clotfelter et al. (2006) examined the relationship between teaching experience and student achievement. They found that teaching experience has a positive effect on student achievement. The impact was large for math but smaller in reading.

Aaronson et al. (2007) examined teacher experience and student achievement in Chicago public schools. The study used a gain score approach with controls for student and teacher fixed effects. The results showed a strong relationship between teacher characteristics and student achievement. However, traditional measures of teacher quality like teaching experience had little impact on classroom results.

Buddin and Zamarro (2008) used a valued added approach to examine teacher characteristics and student achievement. One of the teacher characteristics examined was teaching experience. Buddin and Zamarro (2008) found that teacher experience is positively related to student achievement but the linkage is weak.

Jacob and Lefgren (2008) examined how differences in teacher quality affect student achievement in a midsized school district. They found small effects for teaching experience. Principal rankings of teachers appeared to be a better predictor of teacher performance than other observable teacher characteristics.

Inexperienced teachers with three years or less of experience do appear to be typically less effective than more experienced teachers. However, after about five years, the differences between the two seem to level off (Rosenholtz, 1986). This may be due to the fact that more experienced teachers do not always continue to seek professional development opportunities and thus stop “growing” in the profession. As beginning teachers continue to grow and learn, they could quickly “catch up” to the level of knowledge of a more experienced teacher.

Type of Certification

Traditionally, federal and state governments have regulated teacher quality with certification and license requirements. In order to gain legal permission to teach, candidates are required to complete an approved education program, pass a state license exam and receive a teacher license. A traditional teacher preparation program requires a bachelor’s degree. The degree program includes courses on subjects as well as methodology on how to teach the subjects. During the course of the degree program, preservice teachers participate in field experiences designed to give the students an opportunity to observe and work with students, teachers and the curriculum in an authentic classroom setting.

Students also participate in clinical experiences during the course of the degree program. Clinical experiences are more tightly controlled educational settings. The

clinical setting is more than observing. This experience provides an opportunity for the preservice teacher to work with teachers, students and the curriculum in a more hands-on fashion and usually while studying a particular component of teaching such as classroom management, lesson design or students with disabilities.

The field experience and clinical experience lead up to the student teaching experience. This experience provides an opportunity for the preservice teacher to gradually assume total teaching responsibility under the joint supervision of a principal, cooperating teacher, and a university supervisor. The cooperating teacher remains the teacher of record and is ultimately responsible.

As a result of the teacher shortage, many districts have turned to alternate pathways to licensure to fill empty classrooms. These pathways do not include the same level of education methodology coursework, field experiences, clinical experiences or student teaching. The qualifications of an alternatively licensed teacher vary from state to state. Generally, alternatively licensed teachers are required to have a bachelor's degree, pass a state licensing exam, participate in local teacher training and mentoring, and begin a teacher education program at an institution of higher education.

At this point, it is important to define the different types of teacher licenses that will be mentioned below. A provisional teaching license is a three year, non-renewable license created to allow individuals with no previous education coursework the opportunity to teach while working to obtain a standard teaching certificate. An emergency teaching license is usually issued to a person who holds a standard teaching license but is being asked to teach out of grade or out of subject. A probationary

license is only issued one time to individuals who have not completed the re-education requirements for the five year or re-entry license.

There are several federal programs designed to recruit candidates for teaching and provide a non-traditional pathway into the classroom. One program is Teach for America. Teach for America is a non-profit organization that recruits college graduates to teach in districts needing teachers. Candidates must have a college degree and participate in teacher training and a mentoring program. The candidates commit to teaching for two years. These teachers are considered to be lateral entry teachers and have a provisional license.

Another alternative teacher licensure program is Troops to Teachers. The Troops to Teachers program connects military personnel who are ready to change careers with school districts. The federal program covers the cost of a teacher preparation program and pays bonuses to those who teach in schools with a high percentage of disadvantaged students. Close to four thousand troops were hired between 2002 and 2004. Over 80% of the troops were male and over 25% were African American. This gender and racial diversity has been considered a strength of the program (Shaul, 2006). These teachers are also considered to be lateral entry teachers and have a provisional license.

Many alternatively licensed teachers come into the profession through the lateral entry program. The requirements of the lateral entry program are similar to those of the Teach for America program. Candidates must have a degree, participate in a training program and a mentoring program, and enroll in a teacher education program.

About 1/3 of all new teachers nationwide enter the classroom through a non-traditional pathway. That amounts to 35,000 classrooms across the nation. This large number has caused educators to consider if the pathway to the classroom is an appropriate measure of teacher quality. Several researchers have attempted to determine if there is a connection between traditional pathways and non-traditional pathways and student achievement. Below is a review of research on the Teach for America program, Troops to Teachers program, and lateral entry program.

Evertson, Hawley, and Zlotnik (1985) reported a consistent positive effect of teachers' formal education training on supervisory ratings and student learning. Eleven of 13 studies showed greater effectiveness for fully prepared and standard licensed teachers vs. non-licensed or provisionally licensed teachers.

Fetler (1999) studied math teachers and students in California. He found that lower math scores were linked to teachers on emergency provisional permits and higher math scores were linked to traditional teacher qualifications and increased years of teaching experience.

Decker, Mayer and Glazerman (2004) attempted to determine if Teach for America teachers improved student outcomes compared to traditionally licensed teachers. To answer this question, they compared the outcomes of students taught by Teach for America teachers and students taught by traditionally licensed teachers in the same schools and at the same grades. Students were randomly assigned to make sure that both groups had similar classes of students. Decker et al. (2004) found that Teach for America teachers had a more positive impact on the math achievement of their students than traditionally licensed teachers. Student achievement outcomes for

reading were the same for Teach for America teachers and traditionally licensed teachers.

Goldhaber and Brewer (2000) conducted a regression analysis of 3,786 12th grade math students and 2,524 science students to determine if there was a difference in achievement of students taught by teachers with probationary, emergency, private school license or no license and students taught by teachers who were traditionally licensed. They found that students who were taught by a teacher who was traditionally licensed in math performed better in math than students who were taught by a teacher with no license. The data was not the same for science. They found that there was no difference in the student achievement of students taught by teachers with emergency, probationary or private school license and traditionally licensed teachers.

Darling-Hammond (2000) critiqued Goldhaber and Brewer's (2000) study and used a larger sample of the National Education Longitudinal Study data to conduct the same analysis. Darling-Hammond (2000) found that traditionally licensed teachers had a greater impact on student achievement in math and science than alternately licensed teachers.

Kerr and Berliner, 2002 conducted an ex-post-facto archival design to study the performance of students in the classes of under-licensed and certified teachers. The term under licensed refers to a teacher who has a teaching license, but not necessarily in the subject or grade in which they teach. Districts provided descriptive information about the teachers as well as their class achievement means. This information was compared to the Arizona Department of Education data. A one way analysis of variance revealed that the reading, math and language scores of students

taught by under-licensed teachers (emergency, temporary and provisional) were not statistically different from one another. Students taught by traditionally licensed teachers outperformed students taught by under-licensed teachers, particularly in reading and language. The difference between licensed and under-licensed teachers was statistically insignificant in math.

Raymond, Fletcher and Luque (2001) examined teacher performance using student and teacher data for the Houston Independent School District from 1996 – 2000. Teach for America teachers were compared to new teachers and experienced teachers in the district. The study examined two aspects of teaching and student achievement. First, they looked at the average performance of Teach for America teachers to how they compared to new and experienced teachers. They also compared the best and worst Teach for America teachers with the best and worst new and experienced teachers. They found that Teach for America teachers had a positive impact on student achievement. They also found that the difference between Teach for America teachers and new and experienced teachers was not statistically significant.

Darling-Hammond, Holtzman, Gatlin, and Vasquez-Heilig (2005) replicated this study using the Houston Independent School District data. They conducted a series of regression analyses on a set of 4th and 5th grade student achievement gains on six different reading and math tests over a six year period. Their study included 271,015 students and 15,344 teachers. For outcome measures, they used student test scores in math and reading on three separate standardized tests. Their findings were similar to Raymond et al. (2001). They found that Teach for America teachers had a positive effect on achievement in math and a non-significant effect in reading.

However, when looking at data over years, they found that Teach for America teachers had a significant negative effect on student scores in math. The effects also became significant and negative in reading over time.

Owings, Kaplan, Nunnery, Marzano, and Blackburn (2006) conducted a national survey of administrators who had Troops to Teachers in their schools. The administrators were asked to evaluate the effectiveness of Troops to Teachers compared to traditionally certified teachers. More than 90% of the administrators reported that Troops to Teachers were more effective in both classroom instruction and classroom management and had a more positive impact on student achievement than traditionally certified teachers.

Kane, Rockoff and Staiger (2006) used six years of student achievement data to examine the effectiveness of licensed, alternatively licensed and non-licensed teachers in the New York public school system. Their results showed that license status has a small impact on student achievement. They determined that teacher experience was a better measure of teacher quality than teacher licensure.

Some teachers also have what is considered to be “advanced certification”. In the late 1980’s the National Board for Professional Teaching Standards created an advanced professional certification system for teachers. National Board Certification is a voluntary, standards-based advanced teaching credential that goes beyond state licensure. Candidates must hold a bachelor’s degree, have completed three full years of teaching, and possess a valid state teaching license. This was an attempt to professionalize teaching in a way that would attract and retain high quality teachers. National Board certification indicates experience and advanced professional

development. Since its inception, researchers have considered whether or not National Board Certification has a positive impact on student achievement and whether or not it is an accurate measure of teacher quality.

The first empirical report to study the effectiveness of National Board Certified teachers was Stone's (2002) mixed model statistical analysis of the classroom effectiveness of National Board Certified teachers. Data was available in Tennessee's Value Added Assessment System. Stone (2002) studied 123 teachers by subject and by year achievement gain scores. Eighteen of the National Board Certified teachers received scores that ranked them as exceptional and 13 were considered to be substantially below average. Stone (2002) also found that none of the teachers produced the achievement gains necessary to earn the status of exceptional according to the state's criterion.

Goldhaber and Anthony (2004) compared 303 of North Carolina's National Board Certified teachers to their non-Board certified peers on student achievement in 3rd, 4th, and 5th grade reading and math. They studied nearly 400,000 students. Their results showed that students taught by National Board Certified teachers gained slightly more than those taught by non-Board certified teachers. Goldhaber and Anthony (2004) noted that the differences were so small that National Board Certification alone will not make the gains required to close the achievement gap between on grade level and below grade level students.

Vandevoort, Amrein-Beardsley, and Berliner (2004) solicited all Arizona National Board Certified teachers for their student tests scores and asked them to complete a questionnaire. Thirty five National Board Certified teachers participated in

this study. Vandevoot et al. (2004) used scale scores in reading and math for 1999 - 2003 school years to compare the achievement gains of National Board Certified teachers and non-Board certified teachers. National Board Certified teachers outperformed their non-Board certified peers in 11 of the 48 comparisons. Of the remaining 37 comparisons, there were 13 in which the non-Board certified teachers produced greater gains than their National Board Certified peers.

Cavalluzzo (2004) used a traditional production function formula to examine the performance of students taught by National Board Certified teachers in Miami Dade County Public Schools. Data included FCAT math scores from over 100,000 ninth and tenth grade students. Cavalluzzo (2004) found that National Board Certified teachers had higher post test scores than their non-Board certified peers, but the difference was not statistically significant.

Sanders, Ashton and Wright (2005) conducted a study to measure the effects of National Board Certified Teachers on the quality of teaching and student achievement. They used end of grade math and reading data from two large North Carolina school districts (Charlotte Mecklenburg Schools and Wake County Schools) from 1999-2000 through 2002-2003. Over 260,000 student records from grades 4 -8 represented over 4600 teacher-subject-grade-year combinations. Sanders et al. (2005) found that students of National Board Certified teachers did not have significantly better rates of academic progress than students of other teachers. They also found that there was a great deal of variation among the National Board Certified teachers.

Goldhaber and Anthony (2007) studied the relationship between National Board Certified teachers and the student achievement of elementary students. They

found that National Board Certified teachers are generally more effective than non-Board certified teachers. The statistical significance differs by grade level and student type. There was no evidence that National Board Certification itself has any impact on teacher effectiveness.

Harris and Sass (2007) used a value added model to relate student achievement to determine if National Board Certified teachers are more effective and have a greater impact on student achievement than non-Board certified teachers. After controlling for years of experience and advanced degrees, Harris and Sass (2007) found that National Board Certified teachers were more effective and had higher student achievement scores in reading than non-Board certified teachers. They found no such differences in math.

Cantrell and Kane (2007) studied 99 pairs of teachers teaching in the same school, grade and year. Each pair had one teacher who was National Board Certified and one non-Board certified teacher with at least three years of experience. Students were randomly assigned to classes. They found that there was no statistically significant difference between the student achievement scores of students of National Board Certified Teachers or non-Board certified teachers.

One would assume that standard teacher licensure or advanced teacher certification would produce greater results in student achievement than provisional, probationary or emergency licensure. However, the research above is mixed in support of this premise. Of course there are instances where quality teachers have standard licensure or advanced certification. There are also instances where quality teachers have provisional, probationary or emergency licensure. Licensure status or pathway

into the classroom is not a valid or reliable measure of teacher quality. Preservice teachers need a better measure that is both valid and reliable in identifying and measuring characteristics of teacher quality.

As much as one would expect variables such as general academic ability and intelligence, subject matter knowledge, advanced degrees, teaching experience, and type of certification to be measurable characteristics of teacher quality, research does not consistently support this premise. That is not to say that the demographic variables mentioned above do not support learning in the classroom and in some cases may lead to increased student achievement. However, in this age of accountability, the profession needs a valid and reliable tool to help preservice teachers identify and measure characteristics of teacher quality.

The use of professional teaching standards would be a much better measure of teacher quality than the measures mentioned above. INTASC standards were developed especially for preservice and beginning teachers. Therefore these standards should be used when attempting to measure teacher quality.

Teacher Performance Evaluation

It is interesting that Jacob and Lefgren (2008) considered principal rankings of teachers to be the best indicator of teacher quality. The researchers conducted a study to examine how well principals can distinguish between effective teaching and ineffective teaching. They measured principals' direct observation of teacher performance in the classroom against more traditional measures of teacher quality including education and experience as well as value added measures of teacher quality based on student achievement gains. Jacob and Lefgren (2008) surveyed all principals

in a mid-sized school district in western United States. Principals were asked to rank their teachers on a variety of performance measures based on their observations. The performance measures were: overall teacher effectiveness, dedication and work ethic, organization, classroom management, raising student achievement in math, raising student achievement in reading, student satisfaction with teacher, parent satisfaction with teacher, positive relationship with colleagues, and positive relationship with administrators.

Jacob and Lefgren (2008) found that principals were skilled at identifying teachers who produce the largest and smallest standardized achievement gains in reading and math. Principals were able to identify teachers who fell into the top 10% and bottom 20%. The principals had greater difficulty measuring teachers who fell into the middle of the standardized achievement gains distribution (60% - 80%).

From their research, Jacob and Lefgren (2008) concluded that principals gather their data about teachers from three main sources: formal and informal observations, reports from parents and students and student achievement data. There was a great deal of variation regarding how principals perceive or respond to parent and student reports. There was also a great deal of variation around how principals use student achievement data to assess teachers. There was less variation around how principals rank teachers based on their formal and informal observations.

This supports the premise that an effective way to identify teacher quality is through the use of direct observation of teacher performance in the classroom using agreed upon performance measures. Demographic variables like general academic ability and intelligence, subject matter knowledge, knowledge of teaching and

learning, advanced degree, teaching experience and type of certification to identify teacher quality have proven to provide inconsistent results.

Evaluation of teacher performance is not a new phenomenon. However, as the focus on teacher quality has developed and demographics and teaching contexts have changed, so have direct observation and evaluation of teacher performance. For example, in the early 1900's teachers were largely evaluated based on moral and ethical behaviors. Good teachers were upstanding members of the community, mostly single women with a limited education. These "school marms" were evaluated more on their personal characteristics than a knowledge base of teaching and learning. This evaluation was widely accepted because at that time, researchers were having a difficult time linking specific teacher behavior to student achievement (Medley & Mitzel, 1963; Morsh & Wilder, 1954).

In the 1960's there began a shift from teachers' personal traits to teaching behaviors in the classroom. There was also a shift from general effectiveness ratings to direct measures of student achievement (Flanders, 1970; Soar, 1972; Medley, 1977). During this time, there was also an improvement in research methodology that allowed for better correlational findings.

The 1980's saw a renewed focus on school reform. The two buzzwords of this reform movement were evaluation and accountability. Brophy and Good (1986) reviewed research spanning two decades and concluded that student achievement was improved when teachers used certain teaching methods. "Effective teaching methods" included focus on the academic objective of the lesson, classroom management that supports maximum time for academic learning, teaching at a brisk pace, and adapting

the curriculum to meet the needs of the students. Brophy and Good's (1986) research was influential in creating teacher evaluation systems and instruments that focused on a core set of behaviors and skills deemed necessary for effective teaching.

Observation checklists based on these behaviors and skills were developed and used for classroom observations. Both preservice teachers and in-service teachers were evaluated on these effective teacher criteria.

Performance evaluation programs emerged in the 1980s as a response to a Brophy & Good's (1986) research. While programs varied across the nation, the primary components were teacher observation, professional development and teacher support and mentoring. With Brophy and Good's (1986) teacher effectiveness research, it became even more evident that in order to improve student learning, reformers must improve teaching. It also became evident that teacher observation instruments measuring specific teaching methods were a better way to evaluate teacher effectiveness. Teacher performance evaluation programs of the 1980's took the teacher observation even further by using the data to identify strengths and weaknesses and to provide opportunities for professional development and teacher support and mentoring.

There were critics of the developing teacher evaluation programs. While many understood that the early evaluation efforts were birthed from effective teaching research, they were concerned about other factors that could impact teaching in the classroom; for example student characteristics, subject matter demands, and instruction goals. There were also concerns that the early teacher observation instruments were highly technical and created a one size fits all evaluation system.

The emerging teacher evaluation programs were questioned by researchers regarding their appropriateness for the profession. Teaching is a highly sophisticated and complex exercise that is both an art and a science. How does one measure

teaching that responds to human diversity and aims for cognitive flexibility... using a wide range of teaching strategies that are activated by sophisticated judgments grounded in disciplined experimentation, insightful interpretation of (often ambiguous) events, and continuous reflection. This kind of teaching aims to diagnose and make use of variability, rather than implement uniform techniques or routines (Darling-Hammond & Snyder, 2000, p.526).

Despite the criticism, educators could not deny the value of evaluation programs based on what teachers know and are able to do in the classroom. This type of evaluation clearly addressed the issue of teacher accountability. For the first time, there was a way to connect teaching methods to student achievement. There were two decades of effective teaching methods research to back that premise. Educators did recognize that perhaps there was a need to broaden teacher evaluation programs to include the concerns of the critics.

Teacher evaluation programs continued to evolve using the effective teaching research until the standards-based movement took hold with National Board Certification in the late 1990s. As teaching standards were developed, they became the new focus of teacher evaluation. These standards did not completely abandon the effective teaching research. Educators continued to be concerned about teachers' understanding of lesson objectives, classroom management, pacing, and meeting the needs of students. New evaluation programs used concepts from effective teaching literature as well as newer concepts of teaching and learning (pedagogical content knowledge). The new aim was to develop teachers who

Possess broad and deep understanding of children, the subjects they teach, the nature of learning and schooling, and the world around them...[and who] exemplify the critical thinking they strive to develop in students (Holmes Group, 1986, p. 28).

Proponents of standards-based evaluation recognize that traditional teacher evaluation programs have been weak on teacher accountability and instructional improvement. They argue that traditional teacher evaluation programs have been based on narrow concepts of teaching, limited opportunities to gather evidence, and only one-way communication. Standards-based teacher evaluation offered a more comprehensive program with clear expectations for performance. New programs also provided opportunities for reflection and two-way communication.

Some states have adopted teacher evaluation approaches that are based on teaching standards such as those developed for National Board Certification and the INTASC Standards for beginning teachers. In particular, states are trying to create teacher performance evaluation programs that not only measure teacher performance, but link teacher performance to improved student achievement. This has not been an easy task. Effective teaching can be difficult to measure; therefore developing appropriate evaluation programs has been traditionally difficult. Not only have the teacher performance evaluation programs been difficult to organize, teacher observation instruments have been difficult to develop and have often been ineffective in improving performance or measuring accountability (Solomon & Podgursky, 2000).

Preservice Teacher Performance Evaluation

Teacher performance evaluation programs, teacher performance observations, professional development, support and mentoring are most critical at the beginning

stages of a teacher's career. Some argue that the most important component of preservice development for teachers is the student teaching experience (Edwards & Briers, 2001). Student teaching provides a safe environment for preservice teachers to face challenges and experience successes as well as failures. Performance observation and feedback is critical at this time. Performance observations provide teachers an opportunity to analyze their own development against standards and expectations for fully licensed teachers. This is when preservice teachers begin to shape their ability to impact student achievement. They develop attitudes about the profession and make decisions regarding their entrance and stay in the profession.

Teacher education programs face the challenge of developing an evaluation program that will provide support and professional development for preservice teachers. The result of this challenge is a large variety of student teaching evaluation approaches that often has little evidence of validity or reliability. Ginsberg and Whaley (2003) surveyed 27 universities with teacher education programs with at least 100 students. They found that the evaluation approaches used by these universities were diverse and included checklists, faculty conferences, student conferences and committee reviews. None of the evaluation approaches were well defined. Most of the teacher evaluation strategies fell into one of four categories: case studies, exhibitions, action research and portfolios.

Many teacher preparation programs use case studies to help preservice teachers examine typical aspects of the classroom experience and to think through their own future behavior. Case studies require students to analyze and interpret narratives about students, teaching events, and teaching and learning environments, but this approach

does not require the preservice teacher to actually teach children in a classroom.

Because there is no direct observation of the preservice teacher, case studies do not evaluate classroom performance based on standards.

Some teacher preparation programs use action research as a way to promote a deep understanding of teaching in context. Preservice teachers are asked to design and conduct investigations related to concerns in the classroom. The questions may be similar to those found in case studies, but the investigation extends beyond reflection to a much more structured study. This activity is designed to remove a sense of isolation and through sharing, moves the act of teaching from a private activity to community property. While this approach increases collaboration and builds a knowledge base that can strengthen teaching, it does not require the preservice teacher to actually teach in the classroom. It also does not provide an opportunity for direct observation of the preservice teacher. It is a creative approach to collaboration but not an evaluation of classroom performance based on standards.

Teacher preparation programs also use exhibitions and portfolios to evaluate and support preservice teachers. These approaches are similar. Exhibitions attempt to analyze the performance of preservice teachers by asking them to directly address a teaching problem. Similar to action research, exhibitions give preservice teachers an opportunity to demonstrate teaching abilities in context. Exhibitions require preservice teachers to answer questions through the presentation of artifacts including observations, lesson plans and videotapes. Portfolios allow preservice teachers to assemble artifacts as well. These artifacts represent a comprehensive and holistic examination of their abilities. Artifacts are collected over time and come from multiple

sources and diverse contexts. The artifacts provide evidence of the preservice teacher's thinking, learning and performance. Exhibitions and portfolios provide an opportunity for preservice teachers to culminate their teaching experience through artifacts that can be reflected upon throughout the beginning of their career. However, these approaches do not require preservice teachers to teach. Without an opportunity for direct observation of the preservice teacher in the classroom setting, these approaches do not represent an evaluation of classroom performance based on standards.

These four categories represent a creative way for student teachers to culminate their student teaching experience, but they are not observation instruments. These evaluative approaches do not require preservice teachers to teach in the classroom. Furthermore, they do not include direct observation of preservice teachers in the classroom. They do not provide an opportunity to measure what preservice teachers know and are able to do in the classroom. It is essential in the field of education to have a preservice teacher evaluation tool that has a sound and defensible conceptual basis based on standards. While the approaches mentioned above could be helpful in problem solving specific issues in education, main indicators of performance are not directly observed and are not standards based.

When specifically searching for literature on direct observation instruments that can be used with preservice teachers, there was a void. There was very little mention of any preservice teacher observation instrument and no mention of a preservice teaching observation instrument based on teaching standards. There was a great deal of discussion on assessment of preservice teachers using the approaches mentioned above (case studies, portfolios, exhibitions, action research).

There is a need for a valid and reliable instrument that can be used for direct observation of what preservice teachers know and are able to do in the classroom setting. There is a need for an instrument that allows for direct observation of preservice teacher performance. There is also a need for a direct observation instrument that clearly defines characteristics of teacher quality through the use of standards. The preservice teacher performance evaluation approaches mentioned above raise more questions than answers. What knowledge is essential for teaching? What is the nature of expertise in teaching? How do we develop valid indicators of teacher quality? On what basis should teachers be evaluated? What can you fairly hold teachers accountable for? How do we decide what teachers should know and be able to do? (Ingvarson & Rowe, 2007) Rowe (2007) believes that these questions can be answered through standards.

Teachers are the most valuable resource available to schools. There is a need in the field of education for investment in future educators during their preservice phase. One of the most critical investments is the use of a direct observation instrument that measures teacher quality through the use of standards in the authentic setting of the classroom. Currently, the field of education is lacking such an instrument.

Summary

Educational reform continues to focus on teacher quality as the primary factor in improving student achievement. Despite the knowledge that teacher quality is important, there continues to be questions related to appropriate measures of teacher quality. Researchers have attempted to measure teacher quality by subject matter knowledge, advanced degrees, teacher test scores, years of experience and level of

teacher certification. Results of this large body of research have been mixed and inconsistent.

Another approach has been to measure teacher quality through teacher performance evaluation programs and teacher performance observation instruments. Long gone are the days of measuring teachers based on their personal values. Extensive effective teacher research has provided a foundation for teacher evaluation programs. Standards-based teacher evaluation programs have improved traditional teacher evaluation programs by providing a more comprehensive approach with clear performance expectations and opportunities for reflection and two-way communication.

Researchers agree that the greatest opportunity to impact teacher effectiveness is during the student teaching experience. Preservice teacher performance evaluation programs can have a long lasting impact on the teaching profession. Performance evaluations can shape teaching styles and orientations and influence commitment and decision to remain in the profession (Rosenholtz, 1986).

This preservice phase provides fertile ground to grow an effective teacher. INTASC standards are designed specifically for preservice and beginning teachers. It is logical to build a preservice teacher performance evaluation program around the INTASC standards. Therefore, it is appropriate to use a preservice teacher observation instrument that measures the INTASC standards. It is important during this phase to provide ongoing feedback and dialogue around strengths, weaknesses, and opportunities for professional development.

The *Student Teaching Assessment Rubric* (STAR) is a preservice evaluation instrument that uses the INTASC standards as its underlying structure. The STAR allows for several opportunities to measure observable INTASC indicators, provides ongoing dialogue between the preservice teacher, cooperating teacher and university supervisor, allows for opportunities to build on strengths and identify weaknesses, identifies areas of professional development and provides ongoing support and mentoring. This study was designed to examine the underlying structure of the STAR.

CHAPTER III: METHODOLOGY AND PROCEDURES

Providing support and evaluative feedback to preservice teachers can have a direct impact on student learning. The purpose of this study was to evaluate the underlying structure of the *Student Teacher Assessment Rubric* (STAR), an instrument developed for evaluating preservice teachers. The primary research question examined in this study was whether the empirical data collected using the STAR is equivalent to the model specification associated with INTASC standards. This chapter will describe the research methodology used in this study. The following sections are included in this chapter: (a) description of population, (b) instrumentation, (c) data collection procedures, (d) data preparation, (e) model specification, (f) model identification, (g) model estimation, (h) model testing, and (i) model modification.

Description of Population and Research Setting

The population for this study was preservice teachers who have completed their educational experience. The subjects in this study were undergraduate and post-baccalaureate students enrolled in a large university in the southeast United States. All of the subjects completed their coursework and were administered the STAR at the end of their student teaching experience.

Archival data were obtained from a large university in the southeastern United States. The university is comprised of seven professional colleges, more than 80 bachelor's degree programs, more than 55 master's degree programs and 12 doctoral

programs. The university serves almost 21,000 (45% men, 55% women) students including more than 4,400 graduate students. The undergraduate population is largely traditional yet diverse including ethnic minorities (25.2%), part-time students, and students with disabilities.

The university's College of Education has 1,313 undergraduate majors and pre-majors and 1,528 graduate students totaling 2,841 students and 93 full-time faculty. The College of Education has been recognized by the State Board of Education and the state's Public School Forum for addressing the teacher shortage using rigorous, innovative graduate-level routes to licensure including Master of Arts in Teaching and "fast track" graduate level initial licensure programs (NCATE Institutional Report, 2005).

Undergraduate students in the College of Education completed their clinical experience by working in a school under the supervision of a cooperating teacher. Students also worked closely with their university supervisor. The STAR was completed during their student teaching experience. The student was observed in the classroom by their cooperating teacher and their university supervisor four times per semester. The cooperating teacher and the university supervisor rated the students using a scale of level 1 to level 4, with level 1 and 2 representing developing performance, level 3 representing performance on target, and level 4 representing exceptional performance. Archival data was used from the 4th and final university supervisor rating.

Instrumentation

The STAR was published in 2007 as a tool designed to assess student teacher performance (Jaus, V.P., Cockman, N. R., Frazier, J. W., Hopper, C. J. & Rebich, S. K., 2007). INTASC standards served as the framework for the development of the STAR. The STAR was designed to provide support and feedback to student teachers in their final semester. The individuals who use the STAR are the cooperating teacher (CT), student teacher (ST), and university supervisor (US). All individuals record their ratings on the same instrument. The instrument is to remain in the student teacher's classroom. There is also an Observation Feedback Form (OFF) that is to be used along with the STAR. The STAR and the OFF are to be used for four observation cycles. The STAR can be used with undergraduate student teachers and graduate student teachers.

The STAR is broken down into ten sections based on the ten INTASC standards. Student teachers are rated on ten standards in four separate agreed upon lesson observations. The first observation evaluates standard 1 (content pedagogy), standard 4 (instructional strategies), standard 5 (motivation and management), standard 7 (planning), and standard 9 (professional growth). The second, third and fourth observations evaluate all 10 standards.

Each standard has two to four elements that are rated. For example, standard 1 (content pedagogy) has three elements that are measured; a. demonstrates knowledge of content, b. implements interdisciplinary approaches and c. multiple perspectives for teaching content, and makes content relevant to learners.

There is an opportunity for the cooperating teacher and the university supervisor to rate the student teacher using levels 1, 2, 3 with 4 being reserved for exceptional/distinguished performance. There is a rubric to help the evaluator determine the level of the preservice teacher being observed. There is also an opportunity to record narrative comments on the OFF.

The purpose of this study was to examine the underlying structure of the STAR. This chapter will provide an overview of the research methodology that was used to answer the research questions. The chapter will include the research methods, setting in which the study was conducted, the procedures that were used, and an overview of the data analysis and a summary.

Research Methods

A confirmatory factor analysis was used to examine the underlying structure of the STAR. A logical sequence of five steps was used to examine the STAR: 1) model specification, 2) model identification, 3) model estimation, 4) model testing, and 5) model modification. The following sections describe the data screening and the five step process.

Procedures

Data Collection

The first step was to seek approval from the Institutional Review Board to use archival data. Upon receiving approval, data was collected from the university. The data was imported from the Excel spreadsheet program into LISREL. The data was screened for outliers, missing data, and the assumptions. Once the data was screened, a

sequence of five logical steps was used to examine the underlying structure of the STAR.

Data screening. Outliers, missing data, and out-of-range values can greatly affect the variance-covariance among variables, and thereby affect the analysis (Schumacker & Lomax, 2004). Therefore, data was screened for outliers, missing data, and out-of-range values. Decisions about deleting outliers were determined by examining the degree of influence each outlier has on the major analysis. The mechanism for missingness was evaluated and imputation methods were utilized. All values that were out-of-range were investigated by going to the original dataset to check for data entry problems.

When using inferential statistics, one is operating under the assumption that the data is distributed normally (Schumacker & Lomax, 2004). Therefore it is important to visually check the data to make sure that it is linear and normal. The assumptions of linearity and multivariate normality were examined. All zero-order bivariate correlations and scatterplots were examined for linearity. Univariate normality was evaluated using skewness coefficients and visual examination of box plots. Multivariate normality was examined using PRELIS. All descriptive statistics and correlation coefficients were reported.

Five Step Process

Model specification. The STAR was designed using the INTASC standards. Each INTASC standard served as a latent variable. There are 10 latent variables: (a) Content Pedagogy, (b) Student Development, (c) Diverse Learners, (d) Instructional Strategies, (e) Motivation and Management, (f) Communication and Technology, (g)

Planning, (h) Assessment, (i) Professional Growth, (j) School and Community Relationships. Table 2 displays the latent variables and the corresponding observed variables. The path between the observed variables and the associated latent variables were estimated. All other paths were fixed. In other words, all observed variables had one path estimated to the *a priori* hypothesized latent variable. It was hypothesized that the theoretical covariance matrix would be equal to the manifest covariance matrix. Furthermore, it was hypothesized that the parameter estimates would be positive and statistically significant.

Table 2

Latent and Observed Variables

<u>Latent Variable</u>	<u>Observed Variable</u>
Content Pedagogy	I1_1 Demonstrates Knowledge of Content I1_2 Makes Content Relevant to Learners I1_3 Implements Interdisciplinary Approaches
Student Development	I2_1 Demonstrates Understanding of Learner Developmental Traits I2_2 Stimulates Reflection to Connect Prior Knowledge to New Concepts I2_3 Provides Opportunities for Student Involvement and Responsibility for Learning
Diverse Learners	I3_1 Individualizes the Instructional Environment I3_2 Meets the Range of Individual Needs I3_3 Sets Expectations for Learning and Achievement
Instructional Strategies	I4_1 Selects Multiple Teaching Strategies I4_2 Utilizes a Variety of Materials and Resources
Motivation/Management	I5_1 Establishes and Maintains a Positive Climate I5_2 Establishes Expectations for Behavior I5_3 Monitors and Responds to Student Behavior I5_4 Manages Time and Materials
Communication/Technology	I6_1 Demonstrates Effective Oral and Written Language I6_2 Poses Quality Questions I6_3 Creates Opportunities for Learner Respons I6_4 Utilizes Media and Technology
Planning	I7_1 Bases Purposeful Learning Activities on

	Essential Skills and District Curriculum
	I7_2 Develops Short and Long-term Planning
	I7_3 Monitors and Adjusts Lesson Plans
Assessment	I8_1 Uses a Variety of Formal and Informal Assessment Strategies
	I8_2 Establishes Criteria and Provides Assessment Feedback
	I8_3 Monitors and Records Assessment Data
Professional Growth	I9_1 Self-evaluates Teaching and the Professional Role
	I9_2 Assumes the Professional Role
	I9_3 Exhibits Leadership Potential within the Classroom, School, and/or Student Teaching Seminar
School and Community Relationships	I10_1 Communicates with Families
	I10-2 Utilizes School and Community Resources

Model identification. To test the model, there must be more values in the sample covariance matrix than number of free parameters being estimated. In this study, there were 915 values in the covariance matrix and 105 numbers of free parameters. This indicated an overidentified model and met the order condition required to test the model. The rank condition was examined by estimating the matrix determinant. It was expected that there will be a nonzero matrix determinant.

Model estimation. Maximum likelihood estimations were used since the data was multivariate normal. If this assumption had been violated, asymptotically distribution free estimation would have been used. The covariance matrix was used as the input file into LISREL.

Model testing. Several fit indices were used to determine the adequacy of the model fit. The criteria used in this study were: (a) nonsignificant chi-square, (b) GFI greater than .89, (c) NFI greater than .89, and (d) RMSEA less than .07 (Schumacker

& Lomax, 2004). Path coefficients were evaluated for statistical significances at the .05 level.

Model modifications. If the initial model did not have an adequate fit, the modification indices were used to determine changes needed to improve the model fit. Only respecification that could be justified was allowed; for example, error variances were allowed to correlate. All modifications were reported and the respecified model was evaluated.

Summary

The purpose of this study was to examine the underlying structure of the STAR. Two research questions guided the study.

- *Question 1-* Does the specified model based on INTASC standards fit the empirical data from the STAR?
- *Question 2-* Are the parameter estimates statistically significant from zero?

An archival data set of responses from university supervisors and cooperating teachers was used. Confirmatory factor analysis was conducted to examine the research questions.

CHAPTER IV: PRESENTATION OF THE DATA

The purpose of this study was to examine the underlying structure of the Student Teacher Assessment Rubric (STAR). A confirmatory factor analysis (CFA) was used to examine the underlying structure of the STAR. The following research questions guided the study: (a) Does the specified model based on INTASC standards fit the empirical data from the STAR? (b) Are the parameter estimates statistically significant from zero? Two CFAs were conducted. The first CFA examined the fit of the data to the model specification. The second CFA was used to cross validate the findings of the first CFA.

Description of Sample

A total of 1264 preservice teachers' ratings on the STAR were included in the archival data set. The data set was divided into two data sets. The first data set consisted of only the cooperating teachers' ratings of the students ($N=575$) and the second data set consisted of the university supervisors' ratings of the preservice teachers ($N=689$). All ratings were collected from 2007 to 2008.

Table 3 provides the frequencies and percentages of students' level (undergraduate or post-baccalaureate), degree program, and area of licensure for each of the data sets. Two levels are represented in the sample; undergraduate students and students returning to the university as a Post-Baccalaureate (Post Baccalaureate) student seeking a teaching license. In the first data set, 145 students (25.22%) were

enrolled in the Post Baccalaureate program to earn a teaching license. Four hundred twenty eight students (74.43%) were seeking an undergraduate degree. In the second data set, 265 students (38.41%) of the students were enrolled in the Post Baccalaureate program to earn a teaching license and 423 students (61.30%) were seeking an undergraduate degree.

Nine degrees were represented: Bachelor of Arts (BA: data set 1 $N=410$ and data set 2 $N= 407$), Bachelor of Fine Arts (BFA: data set 1 $N=2$ and data set 2 $N= 0$), Bachelor of Music (BM: data set 1 $N=8$ and data set 2 $N= 8$), Bachelor of Science (BS: data set 1 $N=1$ and data set 2 $N= 1$), Master of Arts (MA: data set 1 $N=0$ and data set 2 $N= 1$), Master of Arts in Teaching (MAT: data set 1 $N=22$ and data set 2 $N= 50$), Master of Education (MED: data set 1 $N=5$ and data set 2 $N= 7$), Master of Science (MS: data set 1 $N=0$ and data set 2 $N=1$), Doctor of Philosophy (PHD: data set 1 $N=1$ and data set 2 $N= 1$). A few students in each data set were undecided (2, 5). Several students were seeking a teaching certificate ($N=115$, $N=199$).

Ten programs were represented: art education (ARTE: data set 1 $N=17$ and data set 2 $N=17$), child and family development (CHFD: data set 1 $N=42$ and data set 2 $N= 54$), dance (DANC: data set 1 $N=3$ and data set 2 $N= 3$), elementary education (ELED: data set 1 $N=341$ and data set 2 $N= 338$), foreign language (FORL: data set 1 $N=5$ and data set 2 $N= 10$), middle grades (MDLG: data set 1 $N=47$ and data set 2 $N= 73$), music (MUSC: data set 1 $N=9$ and data set 2 $N= 9$), secondary education (SECD: data set 1 $N=62$ and data set 2 $N= 91$), special education (SPED: data set 1 $N=42$ and data set 2 $N= 74$), English as a second language (TESL: data set 1 $N=7$ and data set 2 $N= 10$).

Table 3

Description of Participants Degree Program and Area of Licensure

	Data Set 1		Data Set 2	
	<i>n</i>	%	<i>n</i>	%
Post				
Baccalaureate	145	25.22	265	38.41
Undergrad	428	74.43	423	61.30
BA	410	71.30	407	58.99
BFA	2	.35	0	.00
BM	8	1.39	8	1.16
BS	1	.17	1	.14
CERT	115	20.00	199	28.84
MA	0	.00	1	.14
MAT	22	3.83	50	7.25
MED	5	.87	7	1.01
MS			1	.14
PHD	1	.17	1	.14
UND	2	.35	5	.72
ARTE	17	2.96	17	2.46
CHFD	42	7.30	54	7.83
DANC	3	.52	3	.43
ELED	341	59.30	348	50.43
FORL	5	.87	10	1.45
MDLG	47	8.17	73	10.58
MUSC	9	1.57	9	1.30
SECD	62	10.78	91	13.19
SPED	42	7.30	74	10.72
TESL	7	1.22	10	1.45

Note. The percentages may not add up to 100% due to missing data.

Data Screening

The means and standard deviations of the observed variables for each data set are located in Table 5. There was not a great deal of variability in the scores. For data set 1, the means ranged between 3.14 and 3.61. The standard deviations ranged from .53 to .65. For data set 2, the means ranged from 3.12 to 3.60. The standard deviations

ranged from .48 to .58. There were no significant differences between the two data sets on mean ratings for any of the STAR elements.

Table 4

Means and Standard Deviations for STAR Ratings

	Data Set 1		Data Set 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Demonstrates Knowledge of Content	3.34	.55	3.30	.49
Implements Interdisciplinary Approaches	3.23	.61	3.19	.55
Makes Content Relevant to Learners	3.45	.58	3.41	.54
Demonstrates Understanding of Learner	3.28	.58	3.20	.51
Connect Prior Knowledge to New Concepts	3.39	.56	3.34	.51
Student Involvement and Responsibility	3.33	.54	3.29	.52
Individualizes Instructional Environment	3.29	.63	3.18	.56
Meets the Range of Individual Needs	3.25	.60	3.12	.53
Sets Expectations for Learning	3.49	.58	3.41	.56
Selects Multiple Teaching Strategies	3.31	.62	3.29	.55
Utilizes a Variety of Materials	3.47	.61	3.42	.58
Maintains Positive Climate	3.56	.57	3.53	.54
Establishes Expectations for Behavior	3.38	.59	3.32	.55
Monitors & Responds to Student Behavior	3.31	.61	3.28	.55
Manages Time and Materials	3.39	.59	3.36	.55
Effective Oral and Written Language	3.42	.57	3.33	.54
Poses Quality Questions	3.21	.60	3.12	.54
Opportunities for Learner Response	3.38	.57	3.34	.51
Utilizes Media and Technology	3.32	.63	3.25	.55
Purposeful Learning Activities	3.52	.53	3.48	.52
Develops Short and Long Term Planning	3.27	.62	3.20	.55
Monitors and Adjusts Lesson Plans	3.32	.59	3.24	.53
Formal and Informal Assessment Strategies	3.20	.56	3.13	.48
Provides Assessment Feedback	3.22	.56	3.12	.49
Monitors and Records Assessment Data	3.27	.57	3.17	.49
Self Evaluates Professional Role	3.56	.56	3.50	.56
Assumes Professional Role	3.61	.55	3.60	.52

Exhibits Leadership Potential	3.40	.65	3.38	.58
Communicates with Families	3.14	.63	3.15	.55
Utilizes School and Community Resources	3.29	.63	3.26	.55

Note. Means are based on a 1 to 4 point rating scale.

Before conducting the major analyses, the data were screened for outliers. All values were within acceptable range and no univariate outliers were found. Bivariate scatterplots were examined for linearity and all found to be tenable. All skewness and kurtosis coefficients were less than the absolute value of 1.0 suggesting univariate normality was acceptable.

CFA Results

The underlying structure of the STAR is built on the theory that there are observed variables that are directly tied to specific latent variables. Below are the results of the five-step process used in this study for data set 1.

Five Step Process

Model specification. The STAR was designed using the INTASC standards. Each INTASC standard serves as a latent variable. There are 10 latent variables: (a) Content Pedagogy, (b) Student Development, (c) Diverse Learners, (d) Instructional Strategies, (e) Motivation and Management, (f) Communication and Technology, (g) Planning, (h) Assessment, (i) Professional Growth, (j) School and Community Relationships. Each latent variable has between two and four observed variables. The measurement on the observed variables is intended to provide an overview of the teacher's performance on the INTASC standard which is the latent variable. The study was conducted to determine if this specified model based on the INTASC standards fit

the empirical data from the STAR and to determine if the parameter estimates are statistically significant from zero.

The path between the observed variables and the associated latent variables were estimated. All other paths were fixed. In other words, all observed variables had one path estimated to the *a priori* hypothesized latent variable. It was hypothesized that the theoretical covariance matrix would be equal to the manifest covariance matrix. Furthermore, it was hypothesized that the parameter estimates would be positive and statistically significant.

Model identification. The determinant of the matrix was nonzero and the parameters could be estimated.

Model estimation. Maximum likelihood estimation technique was used. The covariance matrix was used as the input file into LISREL.

Model testing. The estimation of the initial model suggested that the model did not fit as indicated by the following indices: $\chi^2(360, N=575)=3702.91$; $\chi^2/df=10.29$; GFI=.71, NFI=.79, RMSEA = .12. Examination of the correlation coefficients (see Table 5) between the latent variables indicate that there were very high positive values, ranging from .78 to 1.00, suggesting some of the latent variable did not measure distinct factors.

The first six variables were highly correlated at .90 or above (Content Pedagogy, Student Development, Diverse Learners, Instructional Strategies, Motivation and Management, Communication and Technology) indicating that they all measure the same dimension of teaching. The relationship among the last four latent variables were also highly correlated at .80 or above (Planning, Assessment,

Professional Growth, School and Community Relationships) indicating that they measure similar dimensions of teaching.

Table 5

Latent Variables Correlation Matrix

	1	2	3	4	5	6	7	8	9	10
1. Content	1.00									
2. Student	1.00	1.00								
3. Diverse	.94	.96	1.00							
4. Instruct	.95	.96	.95	1.00						
5. Motivate	.90	.90	.88	.90	1.00					
6. Communication	.97	.97	.92	.94	.92	1.00				
7. Planning	.88	.90	.92	.92	.86	.94	1.00			
8. Assessment	.87	.91	.86	.89	.86	.92	.91	1.00		
9. Profession	.86	.84	.82	.86	.87	.88	.86	.81	1.00	
10. School	.82	.82	.79	.79	.78	.83	.80	.80	.80	1.00

Model modifications. The initial model did not have adequate fit according to the fit indices. Because of the extremely high correlation coefficients between many of the latent variables, the number of latent variables was reduced. Based on the high correlations, eight latent variables (i.e., Content Pedagogy to Assessment) were combined into one latent variable and renamed Teaching. The correlation coefficients between these initial eight latent variables tended to be over .90. Two of the originally hypothesized latent variables, Professional Growth and School and Community Relations, were retained as separate factors. A three-factor model was tested and resulted in the following fit statistics: $\chi^2(402, N=575)=5426.69$; $\chi^2/df=13.50$; GFI=.62, NFI=.78, RMSEA = .14. The results suggested that the model did not fit the data. The correlation between the three latent variables ranged from .80 to .89, suggesting that the three factors were highly correlated. A decision was made to test a

one-factor model. Fit statistics for the one-factor model were: χ^2 (405, $N=575$)=5920.87; $\chi^2/df=14.60$; GFI=.60, NFI=.76, RMSEA = .15. The following fit statistics were found after allowing 154 correlated error variances: χ^2 (250, $N=575$)=910.35; $\chi^2/df=3.64$; GFI=.91, NFI=.96, RMSEA = .06. The fit statistics suggested an adequate fit of the data to the one-factor model, but the number of correlated error variances allowed in the model indicated some problems with the model. The path coefficients between the single latent variable and the 30 observed variables were all statistically significant ($p<.01$) and ranged from .77 to .91 (see Table 6).

Table 6

Path Coefficients for Data Set 1

I1_1	.87	I2_3	.86	I4_2	.84	I6_1	.88	I7_2	.87	I9_1	.85
I1_2	.91	I3_1	.89	I5_1	.88	I6_2	.89	I7_3	.88	I9-2	.82
I1_3	.86	I3_2	.88	I5_2	.83	I6_3	.90	I8_1	.86	I9_3	.84
I2_1	.89	I3_3	.88	I5_3	.86	I6_4	.73	I8_2	.87	I10_1	.77
I2_2	.82	I4_1	.90	I5_4	.88	I7_1	.84	I8_3	.83	I10_2	.78

Model validation. Data set 2 was used to examine the fit of the one-factor model. Fit statistics for the one-factor model were: χ^2 (405, $N=689$)=6712.28; $\chi^2/df=16.57$; GFI=.61, NFI=.72, RMSEA = .15. After allowing the 187 error variances to correlate, the fit statistics were: χ^2 (218, $N=689$)=855.89; $\chi^2/df=3.92$; GFI=.92, NFI=.97, RMSEA = .06. These results are similar to the one-factor model using data set 1 (see Table 7).

Table 7

Path Coefficients for Data Set 2

I1_1	.80	I2_3	.83	I4_2	.81	I6_1	.78	I7_2	.82	I9_1	.80
I1_2	.87	I3_1	.82	I5_1	.86	I6_2	.83	I7_3	.82	I9-2	.82
I1_3	.86	I3_2	.83	I5_2	.81	I6_3	.83	I8_1	.84	I9_3	.81
I2_1	.89	I3_3	.84	I5_3	.81	I6_4	.69	I8_2	.85	I10_1	.76
I2_2	.83	I4_1	.85	I5_4	.85	I7_1	.82	I8_3	.83	I10_2	.79

Summary

The results of the CFA suggested that the original model specification of 10-factors was not an adequate fit to the data. The intercorrelations between the latent variables were extremely high ($>.90$) suggesting a one-factor model was a better fit. CFA results for the one-factor model indicated an adequate fit but the high number of correlated error variances ($N=154$) allowed in the model to improve the fit is problematic. Since the correlated error variances were determined by the modification indices, the model may have improved because it was improving captioning on chance. The model validation results confirmed the findings.

CHAPTER V: FINDINGS, IMPLICATIONS, FUTURE RESEARCH

The purpose of this study was to examine the underlying structure of the STAR. The review of literature revealed that the most significant factor to improving student achievement is to improve teacher quality. The most effective way to improve teacher quality is to identify characteristics of quality teachers, provide appropriate measures of those characteristics, and provide appropriate feedback based on those characteristics. As with other professions, the means of identifying quality, measuring quality and providing feedback regarding quality is through professional standards. The teaching profession is guided by four sets of standards. INTASC standards are designed to be used with preservice teachers.

Despite the fact that there are a number of educational standards in the field of education, there is a lack of evaluation instruments based on standards. In order to improve student learning and to improve teaching, there needs to be a focus on identifying standards, measuring standards and providing feedback that promotes professional growth. This study is an attempt to add to the body of research on effective teacher evaluation using standards.

The STAR is a teacher evaluation instrument built upon the INTASC standards and was designed to be used with preservice teachers during their student teaching phase. The STAR has 10 latent variables (INTASC standards). Each latent variable has between 2 to 4 observed variables. Two research questions guided the study:

- Does the specified model based on INTASC standards fit the empirical data from the STAR?
- Are the parameter estimates statistically significant from zero?

A confirmatory factor analysis was used to answer the research questions.

Archival data from a large university in the Southeastern United States was used in the study. The study included STAR scores from the 4th and final university supervisor rating.

Key Findings

This analysis began using the 10-factor model. Due to high inner correlation between the latent variables, the 10-factor confirmatory factor analysis did not fit the data. The original model specification was not supported. The high correlations between the latent variables suggested a single factor. The fit statistics for the one-factor model suggested an adequate fit but the large number of modifications suggested some problems.

Implications

Before examining the implications of the study, it is important to review the scope of the research. This study is limited to preservice teachers and does not provide information regarding teachers who are currently active in the field. Also, only the factors associated with the INTASC standards were evaluated. Lastly, this study was limited to STAR data at one university during the final stage of the student teaching experience. It is also important to consider the limitations to the study that are inherent to the methodology. Data used in this study came from three cohorts of university students. These factors may influence the generalization of the results.

The results do not support using observed variables to measure performance on specific latent variables. According to the data, performance in one area was highly related to performance in all other areas. This may suggest that the STAR measures a global measure of teaching ability.

Perhaps this is the case due to the complexity of teaching. It is very difficult to isolate and measure multiple dimensions of teaching. According to Darling-Hammond & Snyder, (2000), teachers must be able to manage student characteristics, subject matter demands and instructional goals simultaneously. Teachers must be able to respond to human diversity and allow for cognitive flexibility. They must be aware of a wide range of teaching strategies and make sophisticated judgments that are grounded in research. They must also be able to interpret vague, sometimes confusing responses to the lesson and continually reflect on their actions. Teaching is a highly complex exercise that is both an art as well as a science. It may be that the INTASC standards and their elements are too discrete to individually measure.

Another explanation might be that strength in one area fosters strength in another area. For example, strength in Instructional Strategies would lead to strengths in Motivation and Management, Planning, Diverse Learners, and other factors of teaching. Also, weaknesses in one area could lead to weaknesses in another area as in the example above.

One should also consider the possibility of rater bias. University supervisors and cooperating teachers could have a tendency toward a predetermined level for all measures, especially towards the end of the student teaching experience. Therefore, they could artificially inflate their ratings.

One way to improve the STAR would be to increase the number of items that measure each factor. Currently there are two to four items for each latent variable. By adding more items, the quality of the measurement of a single factor would be improved. Of course, adding items would increase the work load for evaluators and may not be feasible.

It is also important to keep in mind that the STAR is meant to be used as a snapshot at one point in time. The STAR is used in a very complex way to measure very complex dimensions of teaching. It is possible to break each latent variable down so much that the purpose of the tool is lost. Based on this data, this instrument should not be used to measure teaching at specific factor levels. Rather, this instrument should be used to provide an overview of performance and to give feedback to improve growth as a teacher.

The results of this study do not imply that the STAR should not be used for providing feedback to preservice teachers. The STAR provides clear guidelines that require multiple ratings across the student teaching experience. It is recommended that the STAR be administered four times during the student teaching experience. There are opportunities for the student teacher, cooperating teacher and university supervisor to collaborate before and after the observation. The student teacher has an opportunity to learn about strengths and weaknesses and to make improvements from one observation to the next. There is a comprehensive rubric that is used when rating the student teacher. There is also a scoring profile with descriptions of the performance levels. This is provided to the student teacher, cooperating teacher and university supervisor. There is an opportunity for the teacher to conduct a self evaluation. This

self evaluation gives the student teacher an opportunity to reflect professionally on their performance in the classroom. Professional reflection is a hallmark of teacher quality.

The cooperating teacher and university supervisor have the opportunity to provide narrative feedback on the Observation Feedback Form (OFF). The student teacher receives this feedback as well as a numerical performance score. Feedback is provided during the post-observation conference. This feedback is important for professional development and growth. The information provided on the OFF is the springboard to much deeper conversations regarding teacher quality.

The STAR fulfills all three components of effective evaluation that promotes teacher quality. It identifies characteristics of teacher quality based on standards (INTASC standards). It provides opportunity to measure those characteristics. It also provides opportunity to provide feedback based on the measurements.

Future Research

There are opportunities for future research on the STAR using a larger population of students from more cohorts during different points in their student teaching experience. The data in this study were collected at the end of the student teaching experience. The majority of the preservice teachers in this study performed well in all areas. That may not be the case at a different university or among a different cohort of students to increase the variability in the data. It would be interesting to see the results if the STAR were used with students at different ability levels.

Also, since this data was collected at the end of the student teaching experience, one may argue that it is expected that the students would perform well

after feedback from the cooperating teacher and the university supervisor. There is an opportunity to conduct a similar study with students early in their student teaching experience.

There is also a need for research on preservice teacher evaluation based on standards and professional development. How can an evaluation instrument help a teacher to grow professionally? How can an evaluation instrument improve teacher quality? The STAR has all of the components to promote professional growth, however that was not the focus of this research. Future research focused in that area could provide much needed insight and build on teacher quality research.

Summary

The STAR is a unique evaluation tool for preservice teachers. This study suggested that the STAR is a unidimensional measure of teacher effectiveness instead of the 10-factor model initially proposed. The underlying structure of the STAR model based on the 10 INTASC standards did not fit the empirical data but a one-factor model adequately fit the data. Use of the STAR with preservice teachers has the potential to grow better teachers by working with them during their most professionally fertile phase of student teaching, but using subscale scores may not be appropriate.

REFERENCES

- Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in Chicago Public High Schools. *Journal of Labor Economics*, 25(1), 95 – 136.
- Andrews, J. W., Blackmon, C. R., & Mackey, J. A. (1980). Preservice performance and the National Teacher Examinations. *Phi Delta Kappan*, 61 (5), 358-359.
- Ashton, P., & Crocker, L. (1987). Systematic study of planned variations: The essential focus of teacher education reform. *Journal of Teacher Education*, 2-8.
- Ayers, J. B. & Quall, G. S. (1979). Concurrent and predictive validity of the national teacher examinations. *Journal of Educational Research*, 73.
- Beagle, E. G. (1979). Critical variables in mathematics education: Findings from a survey of the empirical literature. Mathematical Association of America, Washington, DC.; National Council of Teachers of Mathematics, Inc., Reston, VA. (ERIC Document Reproduction Service No. ED171515).
- Betts, J. R., Zau, A. C., & Rice, L. A. (2003). Determinants of student achievement: New evidence from San Diego. San Francisco: Public Policy Institute of California.
- Borko, H., Liston, D. & Whitcomb, J. (2007). Genres of empirical research in teacher education. *Journal of Teacher Education*, 58, 3 – 11.
- Brophy, J. & Good, T. L. (1986). Teacher behavior and student achievement. New York: Macmillan.
- Buddin, R. & Zamarro, G. (2008). Teacher quality, teacher licensure tests, and student achievement. Paper prepared for the Institute of Education Sciences.
- Byrne, C. J. (1983). *Teacher knowledge and teacher effectiveness: A literature review, theoretical analysis and discussion of research strategy*. Paper presented at the meeting of the Northwestern Educational Research Association, Ellenville, NY.
- Cantrell, S. & Kane, T. (2007). National Board Certification and teacher effectiveness: Evidence from a random assignment experiment. Spencer Foundation.
- Cavalluzzo, L. C. (2004). Is National Board Certification an effective signal of teacher quality? (Report No. IPR 11204). Alexandria, VA: The CAN Corporation.

- Clotfelter, C., Ladd, H. & Vigdor, J. (2006). Teacher-student matching and the assessment of teacher effectiveness (NBER Working Paper #11936). Cambridge, MA: National Bureau of Economic Research.
- Cochran-Smith, M. (2006). Taking stock in 2006: Evidence, evidence everywhere. *Journal of Education*, 57:1, 6 – 12.
- Council of Chief State School Officers. (2007). *Interstate New Teacher Assessment and Support Consortium*. Retrieved June 15, 2009, from http://www.ccsso.org/projects/Interstate_New_Teacher_Assessment_and_Support_Consortium/
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis* 8:1.
- Darling-Hammond, L. & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and teacher education*, 16(5 – 6), 523 – 545.
- Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Vasquez-Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives*, 13(42).
- Darling-Hammond, L. (2006). Constructing 21st century teacher education. *Journal of Teacher Education*, 57:3, 300 – 314.
- Decker, P., Mayer, D. & Glazerman, S. (2004). The effects of teach for America on students: Findings from a national evaluation. Mathematical Policy Research, Princeton, New Jersey.
- Druva, C. A., & Anderson, R. D. (1983). Science teacher characteristics by teacher behavior and by student outcome: A meta-analysis of research. *Journal of Research in Science Teaching*, 20 (5), 467-479.
- Edwards, M. & Briers, G. (2001). Cooperating teachers' perceptions of important elements of the student teaching experience: A focus group approach with quantitative follow up. Proceedings of the 27th Annual National Agricultural Education Research Conference.
- Evertson, C. M., Hawley, W. D. & Zlotnik, M. (1985). Making a difference in educational quality through teacher education. *Journal of Teacher Education*, 36:3, 2 – 12.
- Feiman-Nemser, S. & Parker, M.B. (1992). Mentoring in context: A comparison of

two U.S. programs for beginning teachers. National Center for Research on Teacher Learning Special Report. Retrieved from <http://ncrti.msu.edu/http/sreports/spring92.pdf>.

Fenstermacher, G. & Richardson, V. (2005). On making determinations of quality in teaching. *Teachers College Record*, 107(1), 186 – 213.

Ferguson, R.F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation*, 28(2), 465-498.

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.

Ferguson, P., & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education*, 44:11, 55 – 63.

Fetler, M. (1999) “High School Staff Characteristics and Mathematics Test Results,” *Education Policy Analysis Archives*, 7, 9.

Flanders, N. A. (1970). Analyzing teaching behavior. Massachusetts: Addison-Wesley Reading.

Frome, P.; Lasater, B.; Cooney, S. (2005). Well qualified teachers and high quality teaching: Are they the same? (Research Brief). Atlanta, GA, Southern Regional Education Board.

Fuller, E., Carpenter, B. & Fuller, G. (2008). Teacher quality and school improvement in Texas secondary schools. Paper prepared of Association for Texas Professional Educators.

Ginsberg, R. & Whaley, D. (2003). Admission and retention policies in teacher preparation programs: Legal and practical issues. *Teacher Educator*, 38, 169 – 189.

Goldhaber, D. (2007). Everyone’s doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources*, XLII(4), 765-794.

Goldhaber, D. & Anthony, E. (2004). Can teacher quality be effectively assessed? *Human Resources*, XLII:4, 765-794.

Goldhaber, D. & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22(2), 129 – 146.

- Griffin, G. (1987). Forward in G. S. Griffin & S. Millies (eds.), *The first years of teaching*. Chicago. Johnson, S. M. & Kardos, S. (2002).
- Guyton, E. & Farokhi, E. (1987). Relationships among academic performance, basic skills, subject matter knowledge, and teaching skills of teacher education graduates. *Journal of Teacher Education*, 38:5, 37 – 42.
- Haney, W., Madaus, G., & Kreitzer, A. (1987). Charms talismanic: Testing teachers for the improvement of American education. In E.Z. Rothkopf (ED.), *Review of Research in Education*, 14, 169 – 238.
- Hanushek, E., Kain, J., O'Brien, D., & Rivkin, S. (2005). The market for teacher quality. NBER Working Paper No. W11154.
- Hanushek, E. & Kimko, D. (2000). Schooling, labor force quality, and the growth of nations. *American Economic Review*.
<http://ideas.repec.org/a/aea/aecrev/v90y2000i5p1184-1208.html>
- Harris, D. & Sass, T. (2007). The effects of NBPTS-certified teachers on student achievement. National Center for the Analysis of Longitudinal Data.
- Hawk, P., Coble, C. R., & Swanson, M. (1985). Certification: It does matter. *Journal of Teacher Education*, 36 (3), 13-15.
- Haycock, K. (2001). Closing the achievement gap. *Educational Leadership*, 58(6),
- Hellfritzch, A. G. (1945). A factor analysis of teacher abilities. *Journal of Experimental Education*, 14:166 – 169.
- Holmes Group, Inc. (1986). *Tomorrow's teacher: A report of the Holmes Group*. East Lansing, MI: The Holmes Group, Inc. ED 270454.
- Ingvarson, L. & Rowe, K. (2007). Conceptualizing and evaluating teacher quality: Substantive and methodological issues. Paper presented at the Economics of Teacher Quality Conference, Australian National University, 5 February, 2007.
- Jacob, B. & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 26, 101 – 36.
- Jaus, V.P., Cockman, N. R., Frazier, J. W., Hopper, C. J. & Rebich, S. K. (2007). *Student teaching assessment rubric*. Dubuque: Kendall/Hunt Publishing Company.
- Jordan, H. R., Mendro, R. L. & Weerasinghe, D. (1997). Teacher affects on

- longitudinal student achievement. A preliminary report on research on Teacher Effectiveness. Presented at the CREATE Annual Meeting, July 1997, Indianapolis, Indiana: Revised 7/16/97.
- Kane, T., Rockoff, J. & Staiger, D. (2006). What does certification tell us about teacher effectiveness? Evidence from New York City. Working Paper 12155, National Bureau of Economic Research.
- Kaplan, L. & Owings, W. (2001). Teacher quality and student achievement: Recommendations for principals. *NASSP Bulletin* 85(628), 64 – 73.
- Kerr, I. & Berliner, D. (2002). The effectiveness of “Teach for America” and other under-certified teachers on student academic achievement: A case of harmful public policy. *Education Policy Analysis Archives*, 10(37),
- LaDuke, D. V. (1945). The measurement of teaching ability. *Journal of Experimental Education*.
- Loughran, J. (2007). Researching teacher education practices. *Journal of Teacher Education*, 58:1, 12 – 20.
- McArthur, J. (1979). Teacher socialization: The first five years. *Alberta Journal of Educational Research*, 25(4), 264 – 274.
- Medley, D. M. (1977). Teacher competence and teacher effectiveness. A review of process-product research. Washington, D. C.: American Association of Colleges for Teacher Education. (ERIC Document Reproduction Service No. ED143629)
- Medley, D. M. & Mitzel, H. E. (1963). Measuring classroom behavior by systematic observation. Chicago: Rand McNally.
- Morsh, J. E. & Wilder, E. W. (1954). Identifying the effective instructor: A review of the Quantitative Studies. Chanut AFB, II: Air Force Personnel and Training Research Center. (ERIC Document Reproduction Service No. ED044371).
- Monk, D.H. & King, J.K. (1994). Multilevel teacher resource effects on pupil performance in secondary mathematics and science: The case of teacher subject-matter preparation. In R. Ehrenberg (Ed.), *Choices and consequences: Contemporary policy issues in education* (pp.29-58). Ithaca, NY: ILR Press
- Murnane, R. J., & Phillips, B. R. (1981). What do effective teachers of inner-city children have in common? *Social Science Research*, 10, 83–100.
- National Board for Professional Teaching Standards. (2008). Retrieved June 15, 2009, from <http://www.nbpts.org/>

- National Commission on Teaching and America's Future (1996). *What Matters Most: Teaching for America's Future. Report of the National Commission on Teaching & America's Future*. Woodbridge, VA: National Commission on Teaching and America's Future.
- National Council for Accreditation of Teacher Education. (n.d.). Retrieved June 15, 2009 from <http://www.ncate.org/documents/NCATEMission.pdf>
- National Council for Accreditation of Teacher Education and the Utah State Office of Education (2005). NCATE 2005 institutional report. David O. McKay School of Education Brigham Young University.
- Nye, B., Konstantopoulos, S., Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237 – 257.
- The Oyez Project, *Jacobellis v. Ohio*, 378 U.S. 184 (1964) available at: (http://oyez.org/cases/1960-1969/1962/1962_11_2) (last visited Monday, July 13, 2009).
- Owings, W., Kaplan, L., Nunnery, J., Marzano, R., Blackburn, D. (2006). Teacher quality and Troops to Teachers: A national study with implications for principals. *NASSP Bulletin*, 90(2), 102-131.
- Perkes, V. A. (1967). High school science teacher preparation, teaching behavior, and student achievement. *Journal of Research in Science Teaching*, 5, 121 – 126.
- Quirk, T. J., Witten, B. J. & Weinberg, S. F. (1973). Review of studies of the concurrent and predictive validity of the National Teacher Examination. *Review of Educational Research*.
- Raymond, M., Fletcher, S. & Luque, J. (2001). *Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Thomas B. Fordham Foundation.
- Rivkin, S. Hanushek, E. & Kain, J. (2005). Teachers, schools and academic achievement. *Econometrica*, 73(2), 417 – 458.
- Rockoff, J. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94(2), 247 – 252.
- Rosenholtz, S. J. (1986). Workplace conditions that affect teacher quality and commitment: Implications for teacher induction programs. *Elementary School Journal*, 89(4), 421 – 39.

- Rostker, L. E. (1945). The measurement of teaching ability. *Journal of Experimental Education*, 14, 5-51.
- Rowan, B., Chiang, F. S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, 70(4), 256-284.
- Rowe, K. J. (2007). Educational effectiveness: The importance of evidence-based teaching practices for the provision of quality teaching and learning standards. In D. M. McInerney (ED.), *Research on Sociocultural Influences on Motivation and Learning (Volume 7, Standards in Education)*. Greenwich, Conn: Information Age Publishing.
- Sanders, W. L. & Rivers, J. C. (2001). 1996 Teacher quality and equity in educational opportunity: Findings and policy implications. Teacher Quality Conference, Stanford University.
- Sanders, W., Ashton, J. & Wright, A. (2005). Comparison of the effects of NBPTS certified teachers with other teachers on the rate of student academic progress. Study and analysis requested by the National Board for Professional Teaching Standards.
- Schalock, D. (1979). Research on teacher selection. In D. C. Berliner (Ed.), *Review of research in education* (Vol. 7). Washington, DC: American Educational Research Association.
- Schumaker, R. E. & Lomax, R. G. (2004). A beginner's guide to structural equation modeling. New York: Taylor and Francis Group.
- Sclan, E., & Darling-Hammond, L. (1992) Beginning teacher performance evaluation: An overview of state policies. (Trends and Issues Paper No. 7). Washington, D. C. ERIC Clearinghouse on Teacher Education.
- Shaul, M. (2006). Troops to teachers: Program brings more men and minorities to the teaching workforce, but education could improve management to enhance results. Report to Congressional Committees, GAO-06-265.
- Skinner, W. A. (1947). *An investigation of factors useful in predicting teaching ability*. Unpublished master's thesis, University of Manchester.
- Soar, R. (1977). An integration of findings from four studies of teacher effectiveness. Massachusetts: Addison Wesley Publishing Company.
- Soar, R. S., Medley, D. M., & Coker, H. (1983). Teacher evaluation: A critique of currently used methods. *Phi Delta Kappan*, 65:4, 239-246.

- Solomon, L. & Podgursky, M. (2000). The pros and cons of performance-based compensation. Paper. Santa Monica, CA: Miliken Family Foundation.
- Stone, J. (2002). The value-added achievement gains of NBPTS-certified teachers in Tennessee: A brief report. (Report No. ED472132) (ERIC Document Reproduction Service No. ED472132).
- Summers, A. & Wolfe, B. (1975a). Which school resources help learning? Efficiency and equity in Philadelphia public schools. Philadelphia: Federal Reserve Bank of Philadelphia.
- Summers, A. & Wolfe, B. (1975b). Equality of educational opportunity quantified: A production function approach. Philadelphia: Federal Reserve Bank of Philadelphia, Department of Research.
- Summers, A. A., & Wolfe, B. L. (1977). Do schools make a difference? *American Economic Review*, 67, 639–652.
- Title II Higher Education Act. (n.d.). Retrieved June 15, 2009, from <http://www.title2.org>
- Vandevoort, L. G., Amrein-Beardsley, A., & Berliner, D. C. (2004). National Board Certified teachers and their students' achievement. *Education Policy Analysis Archives*, 12(46).
- Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research*, 54: 2, 143-178.
- Wenglinsky, H. (2000). How teaching matters: Bringing the classroom back into discussions of teacher quality (Policy Information Center Report). Princeton, NJ: Educational Testing Service.
- Zeichner, K. (2007). Accumulating knowledge across self studies in teacher education. S:1, 36 – 46.