Teacher Sense of Efficacy for Literacy Instruction and Student Reading

Achievement in Grades Three through Eight

by

JENNIFER M. POGGIO

Submitted to the graduate degree program in Curriculum and Teaching and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Arlene Barry

Chairperson, Arlene Barry, Ph.D.

Diane C. Nielsen

Diane Nielsen, Ph.D.

Barbara S. Bradley

Barbara Bradley, Ph.D.

Karen A. Jorgensen

Karen Jorgenson, Ph.D.

Perry Perkins

Perry Perkins, Ed.D.

Date Defended: April 4, 2012

The Dissertation Committee for Jennifer Poggio

certifies that this is the approved version of the following dissertation:

Teacher Sense of Efficacy for Literacy Instruction and Student Reading

Achievement in Grades Three through Eight

Arlene Barry

Chairperson, Arlene Barry, Ph.D.

Diane C. Nielsen

Diane Nielsen, Ph.D.

Barbara Bradley

Barbara Bradley, Ph.D.

Karen A. Jorgensen

Karen Jorgensen, Ph.D.

Perry Perkins

Perry Perkins, Ed.D.

Date approved: April 4, 2012

ABSTRACT

Teacher Sense of Efficacy for Literacy Instruction and Student Reading Achievement in Grades Three through Eight

This study investigated the effects of teacher sense of efficacy for literacy instruction on reading achievement in grades three through eight. In previous studies, a teacher's sense of efficacy was shown to be correlated to student achievement (Ashton & Webb, 1986; Moore & Esselman, 1992). However, research attempting to uncover the extent to which teacher efficacy relates to student achievement in literacy is sparse. Of the few studies that have been conducted relating student achievement to literacy instruction (Anderson, Green, & Lowen, 1988, Borton, 1991 Armor et al., 1979; Ashton & Webb, 1986, Tschannen-Moran, & Johnson, 2011 & Tranz & Gibson, 1986), a global measure of teacher efficacy was used for data collection. Bandura (2001) highlighted the need for a context specific measure when measuring efficacy, "Scales of perceived self-efficacy must be tailored to the particular domains of functioning that are the object of interest." With the need for a content specific measure for literacy apparent, Tschannen- Moran and Johnson (2011) created and validated a literacy specific measure, The Teacher Sense of Efficacy for Literacy Instruction (TSELI).

For this study, the TSELI was used to measure teachers' sense of efficacy for literacy instruction and the reading test of the Kansas Assessment (2008) was used to measure student reading achievement. Teachers with high efficacy were hypothesized to impact students' reading gains positively. The results indicated that there was a statistically significant relationship between teachers' sense of efficacy for literacy instruction and student achievement gains in reading. This study also hypothesized that as grade level increased TSELI scores would decrease. However, results of this study did not support this hypothesis. Additionally, both

student and teacher level characteristics effects on TSELI scores were examined. Teacher-level variables including highest degree obtained and number of years teaching were found to have a significant effect on TSELI scores. Student-level variables of ethnicity and lunch status were deemed significant.

The results of this study enhance the already extensive research base that exists pertaining to teacher efficacy and achievement. However, this study initiates the research base examining teacher sense of efficacy for literacy instruction and reading achievement using the TSELI measure. Findings and recommendations from this study should be used to inform new investigations relating to teacher efficacy and student achievement in reading. As the construct of efficacy continues to evolve, it is important that teachers, policymakers, and administrators understand its effects.

DEDICATION

I dedicate this work to my boys, Graham and Rhys. May you never fear a challenge always remember that you can accomplish your dreams. Know that I will always be here to support and guide you in whatever those dreams may be.

ACKNOWLEDGEMENTS

This work has been completed because of the support of many individuals. First, I thank God for His love and support through this process. His gentle guidance was there when I needed it most.

I would like to collectively and individually acknowledge and thank the members of my committee. I will forever be grateful for my Committee Chair, Arlene Barry, Ph.D. for her wisdom, guidance and expertise as I completed this dissertation. I would like to thank my Co-Chair, Diane Nielson, Ph.D., who served as my mentor in all things pertaining emergent literacy. I would like to especially thank Dr. Nielson for pushing me, especially early in this process, to stretch and grow as my writing developed. To Barbara Bradley, Ph.D. and Karen Jorgenson, Ph.D. you both provided me with the realism of what completing this degree would entail. Your pragmatism and support guided me in ways you may never know, and for that, I thank you. I would also like to thank Perry Perkins, Ed.D., my Minor chair, for taking this project on. You have been a special part of this committee serving as the devil's advocate, always asking, "Why does it matter". You have a way of forcing one to defend one's thoughts, which is the ultimate goal of this process. Thank you.

To the three district Assessment Coordinators who assisted me with my data collection, I am forever grateful and indebted. I have written countless thank you emails and notes to each of you, but my words cannot thank you enough for your assistance in providing me with such an important component of this dissertation. Without the data, this qualitative dissertation would not have been possible.

I want to express my sincere gratitude to all of the women Ph.D.s that I interact with on a daily basis as a part of my career. There are too many of you to name, but you are out there and

you know who you are! The stories you have shared with me about your process, how you balanced work, children and writing a dissertation, kept me going. You all had a way of checking in on me, holding me accountable, and cheering me on. Any time I wanted to quit, I thought of each of you and what it would feel like telling you I had quit. So, for all of you I am grateful. A special note of thanks to my Godmother and Aunt, Nancy Schreck, Ph.D. I vividly remember you finishing your degree when I was a young child. My parents explained to me that you were a doctor now. However, at that age, I thought you were the kind of doctor who could help me when I was sick. Regardless of how many times my parents tried to explain it to me, I simply couldn't grasp the concept. Obviously, I have grown to understand the meaning of Doctor of Philosophy. You have always been a role model for me and our lives continue to be connected in special ways. I love you.

To my parents, John and Joan Grossman, who raised me to be curious. You have always loved and supported me, even when my goals and desires have not been exactly what you envisioned for me. I love you and thank you. To my in-laws, John and Anne Poggio, your support and words of encouragement, have been motivating throughout this process. And, to both sets of grandparents, your assistance with the boys throughout this time has been vital. I could not have done it without your help and I know the boys benefited from quality time with all of you.

To my boys, Rhys and Graham, I hope and pray that you were little enough while I was writing that you did not realize I was missing at times. I worked hard to finish while you were young so I could be there for the fun events you will have coming up in years to come.

Finally, to my husband, Andrew. No words can express my thanks for your help during this process. You always lifted me up when I was down, served as an academic partner and put up with my crazy antics. You always believed in me, as I do in you. Thank you for your partnership, I love you.

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

CHAPTER ONE

Introduction

There is little debate among researchers, policy makers and educational stakeholders that student achievement is an important issue. According to state-by-state statistics reported to the U.S. Department of Education, more than a quarter of U.S. schools are failing under terms of the *No Child Left Behind Act* (2001). Paul Basken (2006) reported in the *Washington Post* that at least 24,470 schools (27% of the national total) did not meet the federal requirement for adequate yearly progress in 2004-2005. Nationally, fewer than 75% of high-school students graduate, and for African American and Hispanic students, graduation rates sink to 50% or below. In some large, failing high schools, graduation rates are as low as 35%. (The Institute for Student Achievement, 2010). Furthermore, the achievement of students in American schools is falling behind. Peterson (2003) notes that the United States still ranks significantly below other advanced industrial democracies and some developing countries in student academic achievement. Alarming national statistics such as these, coupled with new federal mandates, such as *No Child Left Behind,* have created an era of school reform in America.

Context of the Problem

Despite recent, startling trends and an emphasis on educational transformation, reform efforts in education are not a new phenomenon. In 1953, less than a year after the Russian's launched Sputnik, Congress passed the National Defense Education Act (NDEA). The act was a four-year program that poured billions of dollars into the U.S. education system to help it keep up with its Russian counterpart. Furthermore, in 1983, a landmark report was issued by the National Commission on Excellence in Education. *A Nation at Risk* (1983) documented the characteristics and perceived inadequacies of the public-education system in the United States. Public education was challenged for failing to meet the needs of too many students and for failing to produce the achievement gains that the report's authors deemed vital. As a result, educators were challenged to examine their own practices and policies and to consider how best to support real change and improvement.

Additionally, The Elementary and Secondary Education Act was passed in 1965 as a part of President Lyndon Johnson's "War on Poverty." ESEA emphasized equal access to education and established high standards and accountability. The law authorized federally funded education programs that were administered by the states. In 2002, Congress amended ESEA and reauthorized it as the No Child Left Behind Act (NCLB). NCLB (2001) supported standards-based education reform, based on the belief that high standards and measurable goals could improve individual outcomes in education. The Act required states to assess all students in certain grades on basic skills in order to receive federal funding for schools.

More recently, college and career readiness has been become a new measure of educational excellence at the K-12 level, shifting the goal of K-12 education from simply a diploma to preparation for success at the next level. Whether students attend college or training school, the goal of high school is clear: to prepare graduates for life after high school. Furthermore, the newly written Common Core State Standards (2010) were designed to be vigorous and relevant to the real world, providing students with the knowledge and skills necessary to succeed in college and careers and to compete in the global economy. Regardless of the reform effort or how it may evolve, improved student achievement will remain the common goal.

Factors affecting increased student achievement.

A number of unique variables have been found to influence student achievement, including teacher quality, race, gender, motivation, and socioeconomic status. For the purpose of this study, teacher quality was the most important of these variables. Numerous studies have documented that of all the controllable variables, teacher quality had the greatest effect on student learning (Goldhaber, 2002; Sanders & Rivers, 1997; Sanders, Saxton, & Horn, 1997). Reeves (2004) concluded that the quality of a teacher was almost twice as important as every other variable affecting student achievement. Darling-Hammond (2000) contended that measures of teacher quality in the context of teacher preparation and certification were more strongly related to student achievement than reduced class size, overall spending on education, and teacher salaries. Hanushek, Kain, and Rivkin (1999) identified teacher quality in terms of student performance outcomes. Their research of 400,000 students in 3000 schools revealed that teacher quality was the most important school-related factor influencing student achievement. Therefore, it seems apparent that focusing on teacher quality and ways to improve and maintain high teacher quality, should remain at the forefront of efforts to increase student achievement.

Teacher efficacy.

Repeatedly, one measure of teacher belief that has been related to the characteristics of an effective teacher has been teacher efficacy. Teacher efficacy has been defined as the teacher's belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context (Ross, 1998; Tschannen-Moran, Woolfolk-Hoy, & Hoy, 1998). Teacher efficacy has been associated with a number of teacher behaviors in the classroom. Guskey (1988) found that highly efficacious teachers were willing to try new teaching strategies. Highly efficacious teachers also were found to plan more thoroughly, be more organized (Allinder, 1994), persist through difficult times (Coladaraci, 1992), and work with struggling students longer (Gibson & Dembo, 1984). Evans & Trible (1986) found that teachers with high efficacy demonstrated a greater commitment to teaching.

Teachers' sense of efficacy was a powerful construct related to student outcomes, such as motivation (Midgley, Feldlaufer & Eccles, 1989), achievement (Ashton & Webb, 1986), and students' own sense of efficacy (Anderson, Green & Loewen, 1988). Highly efficacious teachers demonstrate behaviors that resulted in increased student achievement. Enochs, Scharman, and Riggs (1995) found that highly efficacious teachers promoted activity-based and experiential learning. Moreover, such teachers were more humanistic in their classroom management style (Woolfolk & Hoy, 1990), demonstrated more student-centered classrooms (Czerniak & Shriver, 1994), and sought assistance more often from other educational professionals (DeMesuita & Drake, 1994), compared to their less efficacious counterparts. Studying self-efficacy and literacy instruction, Tschannen-Moran and Johnson (2011) found that elementary teachers exhibited higher self-efficacy than their middle-school counterparts.

Bandura (1977) pioneered the concept of self-efficacy beliefs as an evaluation of one's capabilities to attain a desired level of performance in a given endeavor. Self-efficacy beliefs determined how people felt, thought, motivated themselves, and behaved. A strong sense of efficacy increased human accomplishment and personal well-being. In contrast, people who doubted their capabilities shied away from difficult tasks that they viewed as personal threats (Badura, 1997).

The study of teacher efficacy was first investigated more than two decades ago with RAND researchers' evaluation of whether teachers believed they could control the reinforcement of their actions (Armor, Conroy-Oseguera, Cox, King, McDonnell, Pascal, Pauly, & Zellman, 1976). The RAND researchers' work originated on Rotter's (1966) locus of control theory and assumed that student learning and motivation were the relevant reinforcers of teaching action. The RAND studies examined reading instruction among low-income and minority students in an urban setting (Armor et al., 1976). The RAND researchers, in search of variables that would explain differences in the effectiveness of certain teachers and methods, examined the extent to which teachers believed that they could control the reinforcement of their actions. They assessed the extent to which teachers believed they could control student motivation and performance and whether teachers believed environmental factors overwhelmed any power they could exert in schools. The researchers found that teacher self-efficacy was positively related to variations in reading achievement among minority students. Students taught by teachers who believed they could significantly influence students' motivation and learning tended to have higher reading achievement than students whose teachers believed that there was little they could do in light of the impediments to learning posed by the environment. The results of the RAND studies sparked interest in the construct of teachers' self-efficacy beliefs, and in the past three decades, researchers sought to expand and refine the notion of teacher self-efficacy and how it was measured.

Measuring teacher efficacy.

Measuring teacher efficacy has challenged those involved in the field of study. Although several measures were created through the years, construct validity and measurement problems arose. In their seminal work, Tschannen-Moran, Woolfolk Hoy, and Hoy (1998) reviewed teacher efficacy literature and examined the concepts underlying teacher efficacy in order to clarify the construct and to improve how it was measured. For their research, Tschannen-Moran et al. (1998) modeled teacher self-efficacy as the combination of a teacher's perceived competence, the task facing the teacher, and the context in which the task was situated.

Moreover, teacher efficacy has been defined as both context and subject-matter specific. A teacher may vary in his or her feelings of competency, depending on the area of study or the type of students with whom he or she is working. Tschannen-Moran & Johnson (2011) noted that one of the greatest challenges facing the measurement of teachers' self-efficacy beliefs pertained to recognizing the ideal level of specificity for measurement. Scales to measure self-efficacy were developed for use with teachers in specific domains, such as Math (Midgley et al., 1989; Pajares, 1996; Pajares & Graham, 1999; Pajares & Miller, 1994), Science (Riggs & Enock, 1990), Special Education (Egyed & Short, 2006; Meijer & Foster, 1988), and Literacy (Tschannen-Moran & Johnson, 2011).

Statement of the Problem

Educators' confidence in their ability to promote students' learning has proven to be a powerful construct for the past 30 years since the RAND studies of reading instruction among low-income and minority students. Perceived sense of efficacy has been found to be related to teachers' motivation and behavior in the classroom, as well as to students' level of achievement (Armor et al., 1976; Ashton & Webb, 1986; Tschannen-Moran et al., 1998; Ross, 1998).

Concerns regarding the reliability of the two-item RAND scale led researchers to initially develop longer and more comprehensive measures of teacher efficacy. One issue in the measurement of teacher efficacy was the degree of specificity of items in the scales and correspondence between tasks for self-efficacy measurement and outcome task. Most of the measures developed in the wake of the RAND scale assessed efficacy as it applied to teaching in general, as opposed to being content or subject-matter specific (Pintrich & Schunk, 1996; Ross, Cousins, & Gaddalla, 1996). Studies determined that teacher efficacy was a specific rather than a generalized expectancy, finding that within-teacher efficacy beliefs were dependent upon the subject area, type of instructional activity, and characteristics of the class (Raudenbush, Rowan, & Cheong, 1992; Ross et al., 1996; Tournaki & Podell, 2005). Consistent with his view that perceptions of efficacy varied across situations and across the many types of tasks that teachers were required to perform, Bandura (2001) advanced that, "Scales of perceived self-efficacy must be tailored to the particular domains of functioning that are the object of interest" (p. 1).

Heeding the calls that emphasized the importance of content and context specificity in measuring perceptions of efficacy, efforts to develop and validate tailored instruments that provided deeper understanding of subject-specific efficacy beliefs of teachers have been undertaken in recent years. In contrast to the earlier, global measures that viewed perceptions of efficacy as generalized personality traits rather than context-specific judgments, current instruments encompass two dimensions of the construct: an assessment of personal competence and an analysis of the task within particular teaching contexts.

These measurement instruments provided researchers with sensitive tools to investigate teachers' levels of efficacy with regard to the various acts they were asked to perform during instruction in a specific content area. In his review of research, Jerald (2007) highlighted teacher behaviors that were related to a teacher's sense of efficacy. Teachers with a stronger sense of efficacy tended to exhibit greater levels of planning and organization, were more open to new ideas and were more willing to experiment with new methods to better meet the needs of their students, were more persistent and resilient when things did not go smoothly, were less critical of students when they made errors, and were less inclined to refer a difficult student to special education. The practical implications of these findings were that teachers who set high goals, were persistent, and tried another strategy when one approach was found to be ineffective were more likely to have successful students (Shaughnessy, 2004).

While there has been considerable assessment of teacher efficacy and its relationship to various aspects of schooling, research that attempted to uncover the extent to which teacher efficacy related to student achievement in literacy is sparse. Armor et al. (1976); Ashton and Webb (1986); Tracz and Gibson (1986); Anderson, Green, and Loewen (1988); and Borton (1991) have examined the relationship between teacher efficacy and student achievement in literacy. All of these studies, however, utilized global measures of teacher efficacy as opposed to context- and subject-matter specific instrumentation. Given the documented power of the construct of teacher efficacy and the importance of reading as it relates to learning outcomes, research that examines the relationship between the efficacy of reading teachers and student achievement is needed.

Purpose of the Study

The purpose of this study was to investigate the effects of teachers' sense of efficacy for literacy instruction on third- through eighth-grade students' achievement on the reading test of the Kansas Assessment (2008). Higher levels of perceived teacher sense of efficacy in literacy instruction were hypothesized to result in increases in the literacy achievement of students, controlling for student background variables such as prior achievement, gender, ethnicity, and socioeconomic status. The influence of individual teacher characteristics on teachers' sense of efficacy for literacy instruction were also examined, including number of years teaching, education level, gender, and professional development experiences. Additionally, the effects of classroom specific characteristics on teachers' sense of efficacy for literacy instruction, including gender, English language status, special education status, and ethnic make-up.

Research Questions

1. Does a teacher's level of sense of efficacy for literacy instruction decrease as grade levels increase?
2. How do teacher characteristics such as gender, the number of years teaching, education level, and professional development influence teachers' sense of efficacy for literacy instruction?
3. How do classroom-specific characteristics, such as gender, English language learner status, special-education status, free-and-reduced lunch status, and students' ethnicity influence teachers' sense of efficacy for literacy instruction?
4. How does a teacher's sense of efficacy for literacy instruction impact student reading achievement?

Limitations

Several methodological and theoretical limitations must be addressed regarding this study.

1. The study sample consisted of teachers from large districts within the state of Kansas who were identified by the researcher. While drawing participants from large districts was designed to minimize the differences that may have existed between urban/suburban and rural districts, the study sample may not be representative of the general population, thereby limiting the ability to generalize results.
2. Given the voluntary nature of the data gathering method – it may be that the teachers who responded differed in some systematic way related to the variables under study than those who did not.

3. The current study did not control for prior differences in student reading ability, meaning that exhibited differences between levels of independent variables may have been impacted by pre-existing differences among groups. While this potential issue was not addressed statistically, efforts were made to obtain the largest sample possible for the current study. While both the teacher and student samples were sufficient for the analytic procedures employed, some techniques involved the analysis of multiple factors with multiple levels which resulted in imbalanced cell sizes. To combat any potential issues, data were thoroughly checked for violations of assumptions for performing statistical tests prior to all analyses.
4. With regard to the measurement of teacher efficacy, the Teacher Sense of Efficacy for Literacy Instruction (TSELI) is a new tool designed specifically to measure efficacy for literacy instruction. While reliability coefficients and validity evidence have been documented in a single published study (Tschannen-Moran & Johnson, 2011), the survey does not have a significant body of applied research supporting it. Furthermore, this measurement of teacher efficacy relied on self-reported responses to the TSELI. Previous research found that teachers often overestimated their content knowledge and ability, which could affect the validity of the measurement of this construct (Cunningham et. al., 2004; Tschannen-Moran & Johnson, 2011).

Definition of Key Terms

Analysis of Variance (ANOVA) is a statistical analysis tool that separates the total variability found within a data set into two components: random and systematic factors. The random factors do not have any statistical influence on the given data set, while the systematic factors do. The ANOVA test is used to determine the impact independent variables have on the dependent variable in a regression analysis.

Hierarchical Linear Modeling (HLM), also referred to as multilevel modeling, is an analytic technique that is useful to examine data that are nested within one another, such as students within classrooms. HLM controls for the nonindependence of observations that occur due to this nesting, as individuals who belong to a group are likely to be similar to one another, resulting in correlated data. HLM also allows for the examination of the variability within and between individuals and groups as well as their interactions (Raudenbush & Byrk, 2002).

Student Reading Achievement is the act of recognizing words in print and comprehending their collective meaning. For the purpose of this study, student reading achievement is measured by the Kansas Assessment (2008), a multiple choice test.

Teacher Efficacy is a teacher's "judgment of his or her capabilities to bring about desired outcomes of student engagement and learning, even among students who may be difficult or unmotivated" (Tschannen-Moran & Woolfolk Hoy, 2001, p.283).

Teacher Sense of Efficacy for Literacy Instruction (TSELI) is a content-specific tool designed to measure a teacher's sense of efficacy for literacy instruction. Created by Tschannen-Moran and Johnson (2004), the TSELI consists of 22 questions that are answered on a nine-point Likert Scale representing 1-Nothing, 3-Very Little, 5-Some Influence, 7-Quite A Bit, and 9-A Great Deal.

Summary

Despite nearly two decades of policy and investment aimed at closing academic achievement gaps and improving overall student achievement, reading scores on the NAEP 2011 assessment among fourth- and eighth-grade students have barely shifted since 1992. Furthermore, between the 2007, 2009, and 2011 tests, fourth-grade scores were flat, while eighth-grade scores increased just one point, returning to the 2002 level. In the past 50 years, a number of efforts have been put in place by the federal government to increase student achievement. The *National Defense Education Act* (1953) was a four-year program that poured billions of dollars into the U.S. education system to help it keep up with its Russian counterpart. *The Elementary and Secondary Education Act* (ESEA) was passed in 1965 as a part of President Lyndon Johnson's "War on Poverty." *ESEA* emphasized equal access to education and established high standards and accountability. In 1983, *A Nation at Risk* documented the characteristics and perceived inadequacies of the public-education system in the United States. The *ESEA* was amended in 2002 and reauthorized as the *No Child Left Behind Act (NCLB)*, which supported standards-based education reform. More recently, the newly written Common Core State Standards have become the current measure of educational excellence at the K-12 level, with a focus on preparing students to succeed in college and careers and to compete in the global economy.

Of the variables that influence student achievement, teacher quality has been found to have a significant effect on student learning. Therefore, focusing on high teacher quality and ways to improve it must be considered as a way to increase student achievement. Teacher efficacy is a measure of teacher belief that has been related to the characteristics of an effective teacher. Teacher efficacy has been defined as the teacher's belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context (Tschannen-Moran et al., 1998; Ross, 1998).

Bandura (1977) pioneered the concept of self-efficacy beliefs as an evaluation of one's capabilities to attain a desired level of performance in a given endeavor. Self-efficacy beliefs determined how people felt, thought, motivated themselves, and behaved. A strong sense of efficacy increased human accomplishment and personal well-being. In contrast, people who doubted their capabilities shied away from difficult tasks that they viewed as personal threats (Badura, 1982; 1997). The study of teacher efficacy was first investigated more than two decades ago with RAND researchers' evaluation of whether teachers believed they could control the reinforcement of their actions. The RAND studies examined reading instruction among low-income and minority students in an urban setting (Armor et al., 1976). The researchers found that teacher self-efficacy was positively related to variations in reading achievement among minority students. The results of the RAND studies sparked interest in the construct of teachers' self-efficacy beliefs, and in the past three decades, researchers have sought to expand and refine the notion of teacher self-efficacy and how it was measured.

Measuring teacher efficacy challenged those involved in the field of study. Teacher efficacy has been described as being composed of two separate constructs: personal teacher efficacy (PTE) and general teacher efficacy (GTE) (Tschannen-Moran, Woolfok Hoy, & Hoy, 1998). Personal teacher efficacy is the confidence a teacher has in his or her ability to successfully use knowledge, experience, and skills to help students learn. Whereas general teacher efficacy is the teacher's belief that teachers in general successfully can teach students in spite of environmental constraints, such as socioeconomic status, race, and students' individual cognitive and emotion needs (Ashton, Olejnik, Crocker, & McAuliffe,1982). Moreover, teacher efficacy has been defined as both context and subject-matter specific. Tschannen-Moran and Johnson (2011) noted that one of the greatest challenges facing the measurement of teachers' self-efficacy beliefs pertained to recognizing the ideal level of specificity for measurement. Research examining the relationship between teacher efficacy and literacy instruction is sparse. Given the documented power of the construct of teacher efficacy and the importance of literacy instruction as it relates to learning outcomes, research that examines the relationship between the efficacy of reading teachers and student achievement is needed.

CHAPTER TWO

Review of the Literature

Literacy development and achievement can depend on the child, the instruction, environmental factors, or a combination of variables. The purpose of this study was to investigate the effects of individual teacher efficacy for literacy instruction on student achievement in third through eighth grades as measured by the reading test of the Kansas Assessment(2008). Teachers with high efficacy were hypothesized to positively impact students' reading gains. A teacher's sense of efficacy has been consistently recognized as an important characteristic of effective teaching and has been positively correlated to teacher and student outcomes (Tschannen-Moran et al., 1998). Since positive efficacy beliefs are associated with better teaching practices and greater school achievement, it is in the interest of schools and districts to promote teachers' development of strong efficacy beliefs.

This chapter first reviews important aspects of literacy, including the role the federal government and professional organizations have taken recently with regard to literacy instruction. The Kansas Assessment (2008) was used to measure student achievement in this study, thus its development, administration, and other key aspects will be reviewed. This is followed by an in-depth explanation of Bandura's social cognitive theory, which has guided much of the research in the area of teacher efficacy. Individual self-efficacy will be examined, and a review of previous research on individual self-efficacy will be undertaken. A variety of methods for measuring the construct of efficacy will be reviewed, as well as some of the challenges associated with measuring teaching efficacy. Specific attention will be given to The Teacher Sense of Efficacy for Language Instruction Scale (TSELI) (Tschannen & Johnson, 2011). Literature concerning the relationship between teacher efficacy and aspects of school in general, as well as the relationship between teacher efficacy and student achievement in literacy specifically, also will be reviewed.

Literacy

It is widely known that teaching children to read is a primary goal of educators. Early reading skills have been found to be predictors of later reading success. Cunningham and Stanovich (1997) conducted a longitudinal effects of a group of students who were administered a battery of reading tasks in first grade. Ten years later, the same students were administered measures of exposure to print, reading comprehension, vocabulary, and general knowledge. First-grade reading ability was a strong predictor of all of the eleventh-grade outcomes. Furthermore, reviewing standardized test performance, amount of time reading was found to be the strongest predictor of achievement (Guthrie, Wigfield, Metsala, & Cox, 1999). Additionally, NAEP (2011) highlighted that reading for enjoyment correlated with higher achievement scores. Findings such as these have encouraged efforts to improve literacy instruction. A number of reports, reviews, curricula updates, redesigned professional development, and new federal initiatives have stemmed from these efforts. The following highlights the major initiatives over the past decade.

Federal Initiatives.

No child left behind.

Over the past several years, federal initiatives have been formed in an effort to ensure accountability of teachers and schools attempting to close the achievement gap. Most notable among these were the reauthorization of ESEA (Elementary and Secondary Education Act, 1965), or No Child Left Behind (NCLB) (Public Law 107-110). NCLB originally was proposed by the administration of President George W. Bush and signed into law on January 23, 2001.

NCLB supported standards-based reform, which was based on the belief that setting high standards and establishing measureable goals could improve individual outcomes in education. NCLB required states to develop assessments to be given to all students in certain grades if those states were to receive federal funding for schools. The NCLB Act did not propose a set of national achievement standards, rather standards were set and measured by each state. Under NCLB, students' scores on statewide standardized tests monitored schools' performance. Furthermore, with the legislation of NCLB, several important new reading initiatives were added to the Elementary and Secondary Education Act: Reading First, Early Reading First, and Striving Readers.

Reading first. Reading First was an initiative under NCLB targeting students in kindergarten through third grade. Formula grants were used, and the State Education Agencies (SEA) awarded subgrants to eligible Local Education Agencies (LEA) on a competitive basis. Schools funded by Reading First were required to use Scientifically Based Reading Research (SBRR) and to hire and use coaches who assisted teachers in implementing the newest SBRR in their classrooms. Furthermore, coaches analyzed data, collected through the use of researched assessment tools, to drive instruction for individual children in every classroom to ensure that they would learn to read by the end of third grade. Professional development aimed at ensuring that all teachers had the skills they needed to teach the chosen programs effectively was also a goal of the grant (NCLB, 2001)

Early reading first. In an effort to prepare young children to enter kindergarten with the necessary language, cognitive, and early reading skills necessary to prevent reading difficulties, Early Reading First was born. As part of the "Good Start, Grow Smart" initiative, Early Reading First was designed to transform existing early-education programs into centers of excellence that provided high-quality early education to young children, focusing on those from low-income families (NCLB, 2001). Federal funds were awarded competitively to local programs that showed they would increase young children's language and cognitive development by providing high-quality instruction and ongoing professional development based on scientifically based research. Moreover, the program was based on the understanding that literacy was a learned skill, not a biological awakening. Thus, preschool-age children were given cognitive learning opportunities in high-quality language and literature-rich environments to help them achieve the essential knowledge and skills necessary for ideal reading development in kindergarten and beyond. Screening assessments were used to effectively identify preschool-age children who might be at risk for reading failure.

Striving readers. The Striving Readers Comprehensive Literacy (SRCL) formula grant was authorized under NCLB to advance literacy skills for students from birth through grade twelve. Striving Readers' aim was twofold. First, to raise middle- and high-school students' literacy levels in Title 1-eligible schools with significant numbers of students reading below grade-levels, and second, to build a strong, scientific research base for identifying and replicating strategies to improve adolescent literacy skills. The grant required the formation of a SRCL State Literacy Team with expertise in literacy development and education for children from birth through grade twelve. Additionally, grantees were required to develop a comprehensive literacy plan (Striving Readers, 2010).

Common core state standards.

National achievement standards were not a component of NCLB. Rather, states set curricular standards individually and measured them on assessments unique to each state. In an effort to bring state curricula into alignment with each other, the Common Core State Standards

(CCSS) were created (Common Core State Standards Initiative, 2010). The roots of the Common Core State Standards can be traced back to 1980 with the *A Nation at Risk* report. In 1989, at the National Governors Association meeting, a proposal was made to create a set of common national standards. In 1990, the National Educational Goals Panel was established to assess and report state and national progress toward achieving the National Education Goals. In 1996, Achieve Incorporated, a nonprofit subsidiary of the National Governors Association, was launched to help states with standards and assessment and to study how other countries addressed these issues. In 2001, with the reauthorization of ESEA as NCLB, national attention became focused on education issues. With funding from the Gates Foundation in 2009, the Common Core State Standards for both Reading (English Language Arts, ELAs) and Mathematics were released in 2010 (Shanahan, 2011).

The Common Core State Standards were built on what experts considered the best standards work of the states. The goals of the CCSS were: (1) improve educational attainments in the United States, (2) standardize educational opportunities, and (3) focus attention on fewer, higher, better standards and outcomes. Accountability was a major component of the CCSS, and the Department of Education has pledged $350 million to develop new and consistent assessments. When the CCSS were released in 2010, two consortia, the Smarter Balanced Assessment Consortium (SMARTER) and the Partnership for Assessment of Readiness for College and Career (PARRC), began creating these assessments. States that chose to use the CCSS must select either SMARTER or PARRC. Districts will be assessed over the CCSS in the 2014-15 academic year. Additionally, colleges and universities will use these assessments as part of the admissions process (Common Core, 2010).

The Common Core State Standards for English Language Arts (ELA) focus on rigorous content and application of knowledge through higher-order skills. These standards have been internationally benchmarked in an effort to better prepare students to compete in a global economy. Furthermore, the ELAs were built with college and career readiness in mind and consist of the following areas: Reading, Writing, Speaking and Listening, and Language. The ELAs use an integrated model of literacy and have media requirements blended throughout. Changes for English Language Arts include: (1) an emphasis on expository text, critical reading, and technology; (2) literacy across academic disciplines; (3) importance of text difficulty; and (4) focus on canonical texts, such as Shakespeare (Shanahan, 2011). Common Core State Standards have not affected this study, but are necessary to note in the current educational landscape.

Race to the top.

In July 2009, President Barack Obama and Secretary of Education Arne Duncan introduced an initiative known as *Race to the Top*, a program intended for education reform and innovation in classrooms. Competitive grants totaling more than $10 billion were given to states and districts leading reform efforts. *Race to the Top* highlighted four noteworthy areas: (1) adopting internationally benchmarked standards and assessments that prepared students for success in college and the workplace; (2) recruiting, developing, rewarding, and retaining effective teachers and principals; (3) building data systems that measured student success and informed teachers and principals how they could improve their practices; and (4) turning around the lowest-performing schools. To be considered for a *Race to the Top* grant, states who apply must adopt the Common Core State Standards (Race to the Top, 2011).

Professional Organizations Reports

Preventing reading difficulties in young children.

In their report, *Preventing Reading Difficulties in Young Children,* Snow, Burns, and Griffin (1998) noted that there was a rising demand for literacy. They pointed out that the demands for high literacy were ever increasing in a technical society, creating more significant consequences for those who fall short. Out of concern for the growing number of American children who did not read well enough to meet the demands of an ever-increasing competitive economy, the U.S. Department of Education and the U.S. Department of Health and Human Services called for the National Academy of Sciences to establish a committee to examine the prevention of reading difficulties. This committee studied the effectiveness of interventions for young children who were at risk of having problems learning to read and made recommendations for parents, educators, publishers, and others involved in the care and instruction of youngsters. The report's recommendations highlighted the following areas:

- Literacy instruction in first through third grades
- Literacy development in preschool and kindergarten
- Education and professional development for all involved in literacy instruction
- Teaching reading to speakers of other languages
- Ensuring adequate resources to meet children's needs
- Addressing the needs of children with persistent reading difficulties

The committee concluded that most reading difficulties could be prevented and that much work must be done to assure "aggressive deployment of the information currently available" (Snow et al., 1998, p.13). The committee made recommendations for future research but noted that with the knowledge currently available, higher levels of literacy for larger numbers for American school children should be possible.

National reading panel.

In 1997, Congress directed that a national panel be formed to review and evaluate research on the effectiveness of various approaches for teaching children to read. This panel came to be known as the National Reading Panel (NRP). Building on the work and recommendations of the NRC committee's *Preventing Reading Difficulties in Young Children* report, the NRP identified and summarized research literature relevant to the critical skills, environments, and early developmental interactions necessary for children to acquire beginning reading skills. The NRC Committee did not specifically address how critical reading skills were most effectively taught and what instructional methods, materials, and approaches were most beneficial for students of varying abilities. At the finalization of its review, the National Reading Panel presented a report that offered their conclusions, ready for application in classrooms (National Institute of Child Health and Human Development, 2000).

The National Reading Panel was comprised of 14 individuals who formed subgroups to review phonemic awareness, phonics, fluency, vocabulary, and comprehension. A brief review of the NRP's findings is provided below.

Phonemic awareness. Phonemes are the smallest units making up spoken language. The English language consists of about 41 phonemes that combine to form syllables and words. Phonemic awareness refers to the ability to focus on and manipulate these phonemes in spoken words. Teaching children to manipulate the sounds in language (phonemes) helps them learn to read. The NRP report found that teaching phonemic awareness to children significantly improved their reading when compared to instruction without attention to phonemic awareness.

Specifically, the results of experimental studies led the NRP to conclude that phonemic awareness instruction improved students' phonemic awareness, reading, and spelling (NRP, 2000).

Phonics. Phonics instruction teaches children how letters correspond to sounds and how to apply that knowledge to reading and spelling. Such instruction can be provided through explicit, systematic instruction in a set of pre-specified associations between letters and sounds. Children are taught to use these associations to read, typically in texts with controlled vocabulary. The NRP found that systematic phonics instruction led to significant positive benefits for students in kindergarten through sixth grade and for children with difficulty learning to read. For example, kindergartners who received systematic beginning-phonics instruction exhibited increased reading and spelling ability, and first graders were better able to decode and spell words. The students also showed significant improvement in their ability to understand what they read (NRP, 2000).

Fluency. Reading fluency is known to be a critical variable for reading comprehension. Furthermore, reading practice is believed to improve fluency. When children read out loud with speed, accuracy and proper expression, they comprehend material better than if they read with difficulty and in an inefficient way (Rasinski, 2003). Two approaches have been used to teach reading fluency. The first is guided repeated oral reading, which involves reading passages out loud with systematic and explicit guidance and feedback from a teacher. The second approach, independent silent reading, encourages students to read silently on their own with little guidance or feedback from teachers. The NRP determined that guided repeated oral reading had a significant and positive impact on word recognition, reading fluency, and comprehension for students of all ages. However, the NRP was unable to conclude that independent silent reading, used in isolation, improved reading fluency. Moreover, the NRP indicated that more research was needed to understand the specific influences that independent silent reading practices had on reading fluency (NRP, 2000).

Comprehension. A number of factors are associated with a student’s ability to comprehend text. The NRP noted three common themes in the research on the development of reading-comprehension skills. First, reading comprehension is a complex cognitive process that cannot be understood without a clear description of the role that vocabulary development and vocabulary instruction play in the understanding of what has been read. Second, comprehension is an active process that requires an intentional and thoughtful interaction between the reader and the text. Third, the preparation of teachers to better equip students to develop and apply reading-comprehension strategies to enhance understanding is highly correlated to students' achievement in this area. The NRP concluded that text comprehension was improved when teachers used a combination of reading-comprehension techniques, such as question answering, question generation, and summarization. When students were able to use such techniques successfully, they performed better in recall, answering questions, generating questions, and summarizing texts (NRP, 2000).

Vocabulary. Vocabulary development has long been considered important for reading comprehension, and a number of different approaches are used to teach vocabulary. The NRP concluded that vocabulary should be taught both directly and indirectly. Important to vocabulary instruction was repetition and seeing vocabulary words several times. Learning in rich contexts, incidental learning, and using computer technology all helped children develop larger vocabularies. Using a combination of methods rather than a single teaching method also was found to improve learning (NRP, 2000).

RAND reading study group.

In 1999, the RAND institute was asked by the Department of Education's Office of Educational Research and Improvement (OERI) how OERI might improve the quality and relevance of the education research it funded. As a result, the RAND Reading Study Group (RRSG) was formed consisting of 14 experts in the field of reading. This group set out to propose strategic guidelines for a long-term research and development program supporting the improvement of reading comprehension. The RAND report emphasized three primary domains of reading comprehension: (a) Instruction: What could be done to encourage the development of proficient reading and prevent reading-comprehension difficulties? (b) Teacher preparation and professional development: What could be done to better prepare teachers to deliver effective comprehension instruction? and (c) Assessment: What must be included in an assessment system for reading comprehension? The RRSG concluded that comprehension involved three elements: (a) the reader who was comprehending, (b) the text to be comprehended, and (c) the activity in which the comprehension was a part, defining comprehension as, "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (p. 2).

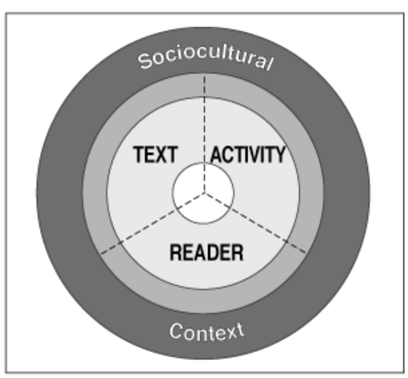


Figure 2.1.
RAND Reading Study-Group Comprehension Heuristic

Following its review, the RRSG made recommendations for further research. The group emphasized the need for research to be practical and useable in classrooms and in guiding educational policy, adding that potential projects must be judged not only by the methodological rigor, but also by their ability to “generate improvements in classroom practices, enhance curricula, enrich teacher preparation, and produce more –informative assessments of reading comprehension” (p. xvi). The importance of using both quantitative and qualitative methodologies, as well as building on existing research when possible, was recommended by the RRSG (RAND, 2002).

Adolescent literacy: A position statement

In 1999, the Commission on Adolescent Literacy, part of the International Reading Association, prepared *Adolescent Literacy: A Position Statement* (Moore, Bean, Birdyshaw, & Rycik, 1999). The aim of this report was to alert school districts and policy makers of key issues related to adolescents’ literacy growth, most notably, the idea that the literacy needs of the adolescent reader differed from those of primary-age children. The Commission outlined seven principles supporting adolescents’ literacy growth:

- Access to a wide variety of reading material
- Instruction that builds both the skill and the desire to read increasingly complex materials
- Assessment that shows students’ strengths and needs and that guides teachers to design instruction to help students grow as readers
- Expert teachers who model and provide explicit instruction in reading comprehension and study strategies across the curriculum
- Reading specialists who assist students having difficulty learning to read and teachers who understand the complexities of individual adolescent readers, respect their differences and respond to their characteristics

- Homes, communities, and a nation that will support efforts to achieve advanced levels of literacy and to provide support for adolescents to succeed (Moore et al., 1999).

Reading next.

Much of the emphasis on improving reading education in the past several years focused on reading in the primary grades. Consequently, the Carnegie Corporation funded the report *Reading Next: A Vision for Action and Research in Middle and High School Literacy* (Biancarosa & Snow, 2004) as part of its 2003 Advancing Adolescent Literacy Initiative. The goal of this initiative was to assist educators in using both research and practice-based knowledge to increase adolescent reading achievement. This report identified 15 elements that were found to be important in improving literacy instruction for struggling readers at the middle- and high-school levels. These 15 elements were split into two categories, instruction and infrastructure. Biancarosa and Snow (2004) noted that, "While the instructional improvements can have a tremendous impact, it is important to realize that they would be more effective if they were implemented in conjunction with infrastructural supports" (p. 12).

Table 2.1
Key Elements in Programs Designed to Improve Adolescent Literacy Achievement in Middle and High School

Instructional Improvements	Infrastructure Improvements
1.Direct, explicit comprehension instruction	10. Extended time for literacy
2. Effective instructional principles embedded in content	11. Professional development
3. Motivation and self-directed learning	12. Ongoing summative assessment of students and programs
4. Text-based collaborative learning	13. Teacher teams
5. Strategic tutoring	14. Leadership
6. Diverse texts	15. A comprehensive and coordinated literacy program
7. Intensive writing	
8. Technology	
9. Ongoing formative assessment	

Summary of Reading

The 2001 reauthorization of the Elementary and Secondary Education Act (ESEA) as No Child Left Behind generated initiatives to increase reading skills at both the elementary- and secondary-education levels. Among them, the Reading First, Early Reading First, and Striving Readers programs were born. A set of national achievement standards was not a component of No Child Left Behind. Consequently, the National Governors Association and the Council for Chief State School Officers worked to develop a set of Common Core State Standards to bring state curricula into alignment with each other. In 2010, Common Core State Standards for English Language Arts and Mathematics were approved. In 2009, President Barack Obama and Secretary of Education Arne Duncan announced a new education initiative: Race to the Top, with the goal of educational reform and innovation in classrooms.

In addition to federal initiatives, a number of reports were commissioned in the past several years to increase reading achievement. *Preventing Reading Difficulties in Young People* (Snow et al., 1998) made recommendations regarding young children at risk of having problems learning to read. In 1997, Congress directed that a national panel be formed to review and evaluate research on the effectiveness of various approaches for teaching children to read. This panel, known as the National Reading Panel, identified and summarized research literature relevant to the critical skills, environments, and early developmental interactions necessary for children to acquire beginning-reading skills including: phonemic awareness, phonics, fluency, vocabulary, and comprehension.

The RAND Reading Study Group (2002) proposed strategic guidelines for a long-term research and development program supporting the improvement of reading comprehension. Emphasizing the need for attention to adolescent literacy, the Commission on Adolescent

Literacy prepared *Adolescent Literacy: A Position Statement* (Moore et al., 1999). The aim of this report was to alert school districts and policy makers of key issues related to adolescents' literacy growth. Furthermore, the Carnegie Corporation funded a report, *Reading Next: A Vision for Action and Research in Middle and High School Literacy* (Biancarosa & Snow, 2004). In this report, Biancarosa and Snow identified 15 elements that were important to improving literacy instruction for struggling readers at the middle- and high-school levels. The elements were split into two categories: instruction and infrastructure.

Effective literacy instruction requires that teachers make multifaceted and immediate teaching decisions to meet their students' needs. Many teachers view teaching students with a wide range of reading levels as one of the greatest challenges they face (Baumann, Hoffman, Duffy-Hester, & Moon, 2000). Effective decision making depends on one's perceived self-efficacy that the knowledge and skills needed to perform the task can be implemented successfully under varied and unpredictable circumstances. Teachers' self-efficacy belief was an important factor in how successful they were. In order to provide higher-quality literacy instruction, teachers' self-efficacy beliefs should be considered.

Self-Efficacy in Education

Teacher self-efficacy

Teacher efficacy has been defined as the extent to which a teacher believes that actions can produce desired outcomes (Ashton & Webb, 1986) as well as teachers' evaluation of their abilities to bring about positive student change (Gibson & Dembo, 1984). More recently, Tschannen-Moran et al., (1998) defined teacher efficacy as "the teacher's belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context" (p. 233). Teacher efficacy, as a broad construct, is the teacher's belief that he or she effectively can influence what students learn.

Tschannen-Moran et al., (1998) reviewed teacher efficacy literature and examined the concepts underlying teacher efficacy in order to clarify the construct and to improve how it was measured. Teacher efficacy has been described as being composed of two constructs: personal teacher efficacy (PTE) and general teacher efficacy (GTE). PTE is the confidence a teacher has in his or her ability to successfully use knowledge, experience, and skills to help students learn. Whereas GTE is the teacher's belief that teachers in general can teach students in spite of environmental constraints, such as socioeconomic status, race, and students' individual cognitive and emotion needs (Ashton et al., 1982). For their research, Tschannen-Moran et al., (1998) modeled teacher self-efficacy as the combination of a teacher's perceived competence, the task facing the teacher, and the context in which the task was situated.

High perception of teacher efficacy has been investigated in relation to a number of teacher behaviors that promote student achievement. In a 1984 study involving classroom observation, Gibson and Dembo found that high-efficacy teachers spent more time in preparation, did not criticize students for giving incorrect answers, were more effective in guiding students to correct answers through skillful questioning, and were more persistent in helping students who were having difficulty. In their study of teachers in middle- and junior-high schools, Ashton and Webb (1986) found that high self-efficacy teachers had less trouble with classroom management; expected students to treat each other and the teacher with respect; were more aware of what was going on in the classroom; kept their students and themselves on task during class; monitored, coached, and praised students more frequently; communicated that school work was important; and rewarded appropriate behavior. On the contrary, Ashton and

Webb found that teachers with low perceptions of efficacy defined their classroom situations in terms of conflict, were more concerned with classroom control than student learning, did not share responsibility for the poor achievement and behavior of low-performing students, and had low expectations of lower-achieving students. This study also related teachers' efficacy beliefs to student achievement; an additional 46% of the variance in student's math achievement was accounted for by adding teachers' sense of efficacy to the regression equation.

Albert bandura's social cognitive theory and teacher self-efficacy.

Original work regarding efficacy stems from psychologist Albert Bandura. Bandura wrote, "Perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the course of action required to produce a given attainment" (p. 3). Bandura (2001) noted that,

> Among the mechanisms of personal agency, none is more central or pervasive than people's belief in their capacity to exercise some measure of control over environmental events…. It is partly on the basis of efficacy beliefs that people choose what challenges to undertake,how much to expend in the endeavor, how long to persevere in the faceof obstacles and failures and whether failures are motivating or demoralizing (p. 10).

The conscious involvement of the individual was central to Bandura's social cognitive theory, rather than behaviorist theories of learning that focused on changing behavior through positive and negative reinforcement. Bandura asserted that behavioral change occurred due to cognitive processes as an individual's cognitive representations of anticipated outcomes assisted in motivating behavior. These cognitive representations created positive outcomes and prevented problems. The reinforcements that resulted from executing the represented action served as the motivation in Bandura's social cognitive theory, rather than simply strengthening an automatic response (Bandura, 1977).

Bandura (1997) stated that control, or agency, was central to human lives, and that predictability allowed people to adaptively prepare for challenges. In contrast, the inability to influence or control factors that may negatively affect one can lead to apathy, apprehension, or despair. Seven effects of efficacy beliefs were central to Bandura's theory: (1) beliefs affected people's choices of actions to take, (2) the amounts of effort they devoted to actions, (3) how persistent they were when faced with challenges, (4) their resilience when faced with adversity, (5) how facilitating or hindering their thought processes were, (6) the amount of stress and depression they experienced in demanding situations, and (7) their levels of achievement. If people did not believe they had the power to produce results, Bandura believed they would not be highly motivated to try.

Bandura's social cognitive self-efficacy theory is often confused with other conceptualizations of self-appraisal. Bandura strove to contrast his theory and make clear his beliefs. Self-efficacy is most often confused with the constructs, self-concept, self-esteem, and locus of control. Self-efficacy is not the same as self-concept, which is defined as "one's collective self-perceptions that are formed through experiences with and interpretations of the environment" (Schunk & Pajares, 2002, p. 16). In a study of college students, Pajares and Miller (1994) found that self-concept in mathematical problem solving did not predict achievement once perceived self-efficacy for mathematical problem solving was factored out. Self-efficacy is also different from self-esteem, which is a judgment of self-worth and self-liking. Self-esteem measures have been shown to be much less predictive of performance than measures of self-efficacy (Mone, Baker, & Jeffires, 1995). Self-efficacy is also quite different from locus of control beliefs (Bandura, 1997). Self-efficacy is context dependent, unlike locus of control, which is a global personal characteristic that applies to many contexts. Also misunderstood is the idea that perceived control is a component of self-efficacy, a more global construct, rather than the same as self-efficacy. People who believe they have some control over what they learn and perform are more likely to invest effort into a learning or performance task (Bandura, 1997). Perceptions of ability, social comparisons, beliefs about the causes of outcomes, available time, and the importance assigned to the task are other factors that influence self-efficacy (Shunk & Pajares, 2002).

Goal setting and self-evaluation, cognitively based motivators, were significant to Bandura's beliefs. Furthermore, Bandura believed that individuals were motivated by standards or performance that they set for themselves and were rewarded by attaining the self-determined level of performance. During self-evaluation, if an individual realized that a discrepancy existed between what he or she hoped to accomplish and what actually was accomplished, the individual was motivated to take action to improve performance. Once the desired level of performance was been achieved, new goals were set for even higher levels of achievement. Bandura asserted that expectations of personal mastery had an effect on both the initiation and the persistence of attempts to reach a certain level of performance (Bandura, 1997).

Bandura suggested that the most daunting elements of human behavior and human motivation were efficacy expectations and outcome expectancy. "Efficacy expectation is a judgment of one's ability to organize and execute given types of performances, whereas outcome expectancy is a judgment of the likely consequences such performances will produce" (Bandura, 1997, p. 21). Furthermore, efficacy beliefs varied in terms of level or magnitude, strength, generality, all of which were major contributors in the structuring of self-efficacy scales used in measurement. According to Bandura, outcome expectations took three forms, which may be described positively or negatively: (a) physical, (b) social, and (c) self-evaluative effects.

As a part of Bandura's (1977, 1997) theory, expectation of personal mastery, or efficacy expectation, was based on four sources of information: (a) enactive mastery experiences, (b) learning through vicarious experiences, (c) verbal persuasion from significant others, and (d) physiological and affective states. Bandura (1997) believed that enactive mastery experiences, or performance accomplishments, were the most important of these. When individuals experienced success in particular situations, expectations were raised for future success when similar situations arose. Enactive mastery experiences provided the individual with the evidence that he or she, "can do whatever it takes to succeed" (p. 80). On the contrary, repeated failures lowered these efficacy expectations. Furthermore, individuals created self-knowledge structures with regard to their efficacy beliefs which were tested each time the individual experienced a situation relating to that efficacy belief. Task difficulty also affected the development of efficacy beliefs. If an individual was faced with a very difficult task, they might be successful, but the effort expenditure and difficulty of undertaking the task actually might lower efficacy beliefs and leave the individual "shaken rather than emboldened" (p. 83).

Learning through vicarious experiences, or modeling, was Bandura's second source of efficacy information. Observing someone else perform a task created a feeling within the individual that they too could achieve at the same levels, because they assessed their capabilities in relation to what others achieved. Models that were judged to be similar in performance capabilities as well as in social and personal characteristics made the most impact on efficacy beliefs (Bandura, 1977; 1997).

The third source of efficacy information was verbal persuasion from a credible other. An individual's sense of efficacy might be boosted through the support of another who provided the individual with the motivation to persist and ultimately succeed in a difficult task. Furthermore, evaluative feedback that focused on an individual's capabilities to achieve a task, especially during skill development, raised efficacy beliefs. However, feedback that focused on the individual's effort expenditure might raise efficacy beliefs in the short term, but might ultimately lead the individual to believe that they succeeded merely through sheer effort rather than actual ability, which could lower efficacy beliefs in the long term. Verbal persuasion was most effective when given shortly after the task was performed successfully. Moreover, the task assigned should only be moderately beyond the individual's current performance level. Likewise, Bandura (1997) noted that efficacy was elevated more when there was not a huge disparity between an individual's own beliefs about his or her capabilities and what they were told about their capabilities by the credible other.

The individual's physiological and affective states served as the final sources of efficacy information. In stressful situations, an individual's heart rate may increase, and they may perspire and have difficulty controlling their breathing. Depending upon the context in which these physiological activities occurred, the individual may relate the causes to different sources. Furthermore, events and circumstances were processed differently when one was sleepy or sad rather than well rested and content, and affective features, such as mood, also could play a role in efficacy beliefs (Bandura, 1997).

Teacher self- efficacy: A content specific construct.

Context also played an important role in the development of teachers' self-efficacy beliefs. Tschannen-Moran et. al., (1998) noted that while assessing their beliefs about teaching capability in a specific context, teachers made two connected judgments: an assessment of their personal teaching competence considering the expected requirements of the teaching task. A teacher made a judgment of personal competence based on an assessment of internal strengths and deficits. Assessment of the teaching task took into consideration such factors as instructional resources available and quality of the curriculum. Student factors considered included student perceived ability, motivation, and socioeconomic status. Contextual variables comprised such variables as school climate, collegial support, and leadership. Additionally, Tschannen-Moran and colleagues (1998) found that teacher efficacy was sensitive to school-level effects, highlighting a reciprocal relationship between school context and teachers' efficacy beliefs. School-wide factors, such as high emphasis among staff on achievement, a strong sense of community in a school, teacher participation in school decision making, receiving positive feedback on teacher performance, collaboration with other teachers, high levels of parental involvement, and school-wide standards for student behavior contributed to higher teacher efficacy.

Career stage also was proven to play a role in teacher efficacy. Tschannen-Moran et. al. (1998) found that as teachers moved on in their careers, efficacy beliefs became less changeable. Moreover, when experienced teachers were required to change practices as a result of district or school initiatives, they experienced an initial decrease in their perceptions of efficacy, however, if they observed an increase in the effectiveness of their new practices, their efficacy would improve greatly. On the contrary, if teachers learned a new strategy and were not successful in implementing it, their self-efficacy declined. Low efficacy beliefs contributed to poor outcomes, which led to lower efficacy beliefs, and the feedback cycle became increasingly negative. Contextual variables such as the availability of teaching resources was found to contribute more to the self-efficacy beliefs of novice teachers than experienced teachers, who had many mastery experiences on which to base their perceptions (Tschannen-Moran & Woolfolk Hoy, 2001). Moreover, verbal persuasion in the form of interpersonal support of administrators, colleagues, parents, and members of the community was shown to greatly impact the variance in novice teachers' self-efficacy beliefs, although not that of experienced teachers. Contextual variables such as school setting were found to be unrelated to the self-efficacy beliefs of both experienced and novice teachers.

Raudenbush, Rowan, and Cheong (1992) found that feelings of efficacy varied between classroom groups, depending on subject content and student characteristics. Furthermore, teacher efficacy was proven to be context and subject matter specific, which resulted in a number of studies relating teacher efficacy to content areas.

Table 2.2
Teacher-efficacy studies done in content areas (including, but not limited to)

Subject	Reference
Math	Midgley, et al, 1989
	Pajares & Miller, 1994
	Pajares, 1996
	Pajares & Graham, 1999
Science	Riggs & Enoch, 1990
Writing	Graham et al., 2010
Classroom Management	Emmer, 1990
Literacy	Tschannen-Moran & Johnson, 2011
Special Education	Egyed & Short, 2006
	Meijer & Foster, 1988

Studies in teachers' perceptions of efficacy.

Teacher efficacy was related to student outcomes, such as motivation (Midgley, Feldlaufer & Eccles, 1989), achievement (Ashton & Webb, 1986), and students' sense of efficacy (Anderson, Green & Loewen, 1988). Moore and Esselman (1992) found that students of teachers with higher levels of efficacy outperformed students of teachers with lower levels of efficacy on the mathematics section of the Iowa Test of Basic Skills. Highly efficacious teachers were willing to try new strategies (Guskey, 1988), persisted when teaching became difficult (Coladaraci, 1992), worked with struggling students longer (Gibson & Dembo, 1984), and were more likely to stay in the profession of teaching (Burley, Hall, Villeme & Brockheimer, 1991). Teacher efficacy and teacher behavior in the classroom also were related. Allinder (1994) found that teachers with high efficacy planned more thoroughly and were more organized. Evans and Tribble (1986) found that teachers with high efficacy showed greater commitment to teaching, while Allinder (1994) reported that high efficacy related to more enthusiasm for teaching.

Ashton and Webb (1986) found that teachers with a high sense of efficacy had high expectations for all students, established classroom environments that encouraged warm interpersonal relationships, and promoted strong academic work. They were more humanistic in their classroom-management style (Woolfolk & Hoy, 1990), evoked trust from their students (DaCosta & Riordan, 1996), and favored student-centered learning (Enochs, Scharman & Riggs, 1995). In addition, teachers with high efficacy were more likely to seek assistance from other educational professionals (DeMusquita & Drake, 1994). Thus, the relationship between teacher efficacy and student achievement was often indirect, as highly efficacious teachers behaved in ways that may result in increased student achievement.

Summary of Self-Efficacy in Education

Teacher efficacy has been defined as “the teacher’s belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context” (Tschannen-Moran et. al., 1998, p. 233). As a broad construct, teacher efficacy is the teacher’s belief that he or she effectively can influence what students learn. Tschannen-Moran et al. (1998) reviewed teacher efficacy literature and examined the concepts underlying teacher efficacy in order to clarify the construct and to improve how it was measured. For their research, Tschannen-Moran et al. modeled teacher self-efficacy as the combination of a teacher's perceived competence, the task facing the teacher, and the context in which the task was situated. High perception of teacher efficacy has been investigated in relation to a number of teacher behaviors that promote student achievement including: more time spent in preparation, not criticizing students for giving incorrect answers, more effective in guiding students to correct answers through skillful questioning, more persistent in helping students who were having difficulty (Gibson & Dembo, 1984), better classroom management, and rewarding appropriate behavior (Ashton & Webb, 1986).

The study of teacher efficacy has its roots in the work of psychologist Albert Bandura. Bandura (1997) wrote, "Perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the course of action required to produce a given attainment" (p. 3). Bandura challenged the beliefs of behaviorist theories, asserting that behavioral change occurred due to cognitive processes as an individual's cognitive representations of anticipated outcomes contributed to motivating behavior. In studies comparing self-efficacy to other conceptualizations for self-appraisal, self-efficacy was shown to be more predictive of student performance (Mone, Baker, & Jeffires, 1995; Schunk & Pajares, 2002). Also key to Bandura's beliefs were efficacy expectations and outcome expectancy. "Efficacy expectation is a judgment of one's ability to organize and execute given types of performances, whereas outcome expectancy is a judgment of the likely consequences such performances will produce" (Bandura, 1997, p. 21). According to Bandura, outcome expectations took three forms: physical, social, and self-evaluative. Furthermore, efficacy expectation (1977, 1997) was based on four sources of information: enactive mastery experiences, learning through vicarious experiences, verbal persuasion from significant others, and physiological and affective states. Teacher efficacy as proven to be context and subject matter specific, which resulted in a number of subject-specific studies related to teacher efficacy. Additionally, Tschannen-Moran et al. (1998) found that teacher efficacy was sensitive to school-level effects, highlighting a reciprocal relationship between school context and teachers' efficacy beliefs.

Methods of Measuring the Construct of Efficacy

Rotter's (1966) social learning theory, Weiner's (1979, 1992) attribution theory, and Bandura's (1977) social cognitive theory have grounded the work of the many attempts to measure teacher efficacy. The RAND organization (Armor et al., 1976) conducted the first studies attempting to measure a teacher's perceptions of his or her own capabilities. The RAND researchers' work, grounded in Rotter's social learning theory (1966), sparked interest in teacher efficacy. Rotter (1966) described perceptions of the outcomes of a person's actions as being controlled by the actor (internal locus of control) or by outside factors (external locus of control). The RAND researchers created and distributed a lengthy questionnaire to measure what they called teacher efficacy. Two questions in the measurement, based upon Rotter's (1966) learning theory, emerged among the most powerful factors.

RAND Item 1. "*When it comes right down to it, a teacher really can't do much, because most of a student's motivation and performance depends on his or her home environment.*"

A teacher who expressed strong agreement with this statement indicated that student outcomes were influenced more by external factors than by teacher actions. Variables outside of school, such as conflict, violence, or substance abuse in the home or community; the value placed on education at home; the social and economic realities of class, race, and gender; and the physiological, emotional, and cognitive needs of a particular child all had a very real impact on a student's motivation and performance in school. Teachers' beliefs about the power of these external factors compared to the influence of teachers and school were labeled general teaching efficacy (GTE) (Ashton & Webb, 1986).

RAND Item 2. "*If I try really hard, I can get through to even the most difficult or unmotivated students.*"

Agreement with this item indicated that the teacher felt confident in his or her ability to overcome factors that would make learning difficult for students (Armor, et al., 1976). These teachers reflected confidence that they had adequate training or experience to develop strategies for overcoming obstacles to student learning and that they had been successful at improving student achievement. This aspect of efficacy was labeled personal teaching efficacy (PTE) (Ashton & Webb, 1986) and was more specific and individual than a belief about what teachers in general could accomplish.

Teachers in the RAND studies were asked to indicate their level of agreement with each of these two statements. The sum of the scores on the two items was called teacher efficacy (TE) and became a construct for measuring the extent to which a teacher believed that he or she controlled the consequences of teaching and student motivation and learning. The RAND study found a significant relationship between teacher efficacy and teachers' success in teaching reading to minority students in urban schools (Armour et al., 1976). In a second RAND study, Berman, McLaughlin, Bass, Pauly, and Zellman (1977) found teacher efficacy to be a strong predictor of the continuation of federally funded projects after the end of funding. Teachers' sense of efficacy had a strong positive effect not only on student performance, but on the percentage of project goals achieved, on the amount of teacher change, and on the continued use of project methods and materials after the project ended.

The results of the RAND studies sparked interest in the construct of teacher efficacy, although the reliability of the two-item scale concerned researchers. Therefore, attempts to develop longer, more comprehensive, measures took place. Two such instruments were Guskey's (1981) Responsibility for Student Achievement (RSA) and Rose and Medway's (1981) Teacher Locus of Control. Both the instruments used Rotter's locus of control as their foundations; conceptualizing teacher efficacy as teachers' beliefs that factors within their control had greater impact on the results of teaching than did factors in the environment or in the student (Tschannen-Moran et al., 1998).

Guskey (1981) developed a 30-item instrument that measured Responsibility for Student Achievement (RSA). For each item on the RSA, participants were asked to distribute 100 percentage points between two alternatives--one stating that the teacher caused the event and the other stating that the event occurred because of factors outside the teacher's immediate control. The following is an example of a scenario from the RSA, "When a student does well in your class, would it probably be because…" A. "the encouragement you offered" or B. "the student had the natural ability to do well." The participant would split the 100 percentage points between the two alternatives.

Guskey framed the alternatives using Weiner's (1979, 1992) attributional theory and created responses that indicated that the occurrence happened due to specific teaching abilities, the effort put into teaching, the task difficulty, or luck. The instrument measured three factors: teacher responsibility for student success (R+), teacher responsibility for student failure (R-), and the combined score of the two, which Guskey named Responsibility for Student Achievement

(RSA). Guskey found a strong correlation between scores on the teacher efficacy scale as developed by the RAND researchers and teacher responsibility for student success (R+) and student failure (R-). He also found strong inter-correlations between overall RSA and responsibility for student success (R+) and student failure (R-). Guskey (1987) asserted that positive and negative performance outcomes represented separate dimensions, not opposite ends of a single continuum, and that these dimensions operated independently in their influence on the perceptions of efficacy. Teachers displayed greater efficacy for positive results than for negative results; teachers were more confident in their ability to influence positive outcomes than to prevent negative ones. Greater efficacy was related to core positive attitudes about teaching, as well as high levels of confidence in teaching abilities on the measure of teaching self-concept (Guskey, 1988).

The Rose and Medway (1981) Teacher Locus of Control (TLC) was a 28-item measure that asked teachers to assign responsibility for student successes or failures by choosing between two competing explanations for the situations described. Half of the items on the TLC described situations of student success, while the other half described student failure. For each success situation, one explanation connected the positive outcome internally to the teacher (I+), while the other assigned responsibility outside the teacher, typically to the students. Likewise, for each failure situation, one explanation gave an internal teacher attribution (I-), while the other blamed external factors. The TLC has been significantly related to the RAND items, although the correlation was weak. Rose and Medway (1981) found that the TLC was a better predictor of teacher behaviors than Rotter’s (1966) Internal-External Scale. For example, the TLC predicted teachers’ willingness to implement new instructional techniques, whereas Rotter’s I-E Scale did not. Teachers who rated high in internal responsibility for student learning in schools with large populations of disadvantaged students gave fewer disciplinary commands. Teachers who rated high in internal responsibility but taught among more privileged students called on non-volunteers more frequently and more often had students engaged in self-directed activities as opposed to listening (Rose & Medway, 1981).

Despite their development and analysis, neither the Guskey nor the Rose and Medway measures were used extensively by other researchers in published literature (Tschannen-Moran & Woolfolk Hoy, 2001). Other measures based on Bandura's (1977) social-cognitive theory also were developed in the early 1980s. The Asthon vignettes, short descriptions of situations teachers might face in their everyday work, were designed by Ashton, Buhr, and Crocker (1984). Teachers were asked to assess how well they thought they would handle each situation and rate it on a Likert scale from "extremely effective" to "extremely ineffective." They then were asked to rate how they would perform in each situation in comparison to other teachers. The comparison question scores were strongly correlated to the two RAND items, but there was no correlation between the RAND items and the self-rating items. Not unlike the instruments grounded in Rotter's work, the Ashton vignettes did not gain wide acceptance. One instrument that did gain acceptance was the Gibson and Dembo Teacher Efficacy Scale (TES) (1984).

Gibson and Dembo's TES (1984) combined concepts from the RAND study and from Bandura's social-cognitive theory. This 30-item measurement asked teachers to rate statements, such as "When a student gets a better grade then she usually gets, it is usually because I have found better ways of teaching," on a six-point Likert scale from strongly disagree to strongly agree. Despite the fact that researchers have questioned the interpretation of the scale (Tschannen-Moran et al., 1998; Woolfolk & Hoy, 1990; Soodak & Podell, 1993), it has been the most widely used measure of teacher efficacy in most literature to date. Ross (1998) stated that almost half of the studies of teacher efficacy conducted up to 1998 used the Gibson and Dembo TES or an adaptation of their scale.

Based on theory developed in their landmark 1998 paper, Tschannen-Moran and Woolfok Hoy (2001) developed a new measure of teacher efficacy: The Teacher Sense of Efficacy Scale (TSES), previously called the Ohio State Teacher Efficacy Scale. They offered a model that combined attributional analysis (internal vs. external) with Bandura's four sources of efficacy information: mastery experience, affect, vicarious experience, and verbal persuasion. Context in teacher efficacy was emphasized by the authors, who stated, "Teachers feel efficacious for teaching particular subjects to certain students in certain settings, and they can be expected to feel more or less efficacious under different circumstances" (Tschannen-Moran et al., 1998, p. 227). Their model presented two factors that were related to, but distinct from, GTE and PTE. They identified the first as the teaching task, and its context reflected teachers' assessment of the difficulty of the task relative to the available resources. A teacher's consideration of the motivation and ability of students, the physical environment in the school, availability of supplies and technology, school leadership, and climate, etc. were included in the first factor. The second factor, self-perceptions of teaching competence, assessed the teacher's "comparative judgment of whether his or her current abilities and strategies are adequate for the teaching task in question" (p. 233).

Research on teachers' self-efficacy for literacy instruction has been at a disadvantage due to the lack of a measure that is content specific, reliable and valid. Szabo and Mokhtari (2004) developed the Reading Teacher Efficacy Instrument (RTEI). This measure was designed based on the Gibson and Dembo (1984) scale and shared many of the same psychometric problems of that measure (Tschannen-Moran & Wolfolk-Hoy, 2001). Tschannen-Moran and Johnson (2011) developed the Teacher Sense of Efficacy for Literacy Instruction Scale (TSELI) to measure teachers' sense of efficacy for literacy instruction so researchers could examine this subject-specific aspect of teachers' self-efficacy beliefs. The authors constructed a set of 33 items specific to various aspects of literacy instruction, drawing on the work of the NCTE and IRA (1996) *Standards for the English Language Arts* and the IRA (2004) *Standards for Reading Professionals.* Items on this instrument measured aspects of literacy instruction, such as the ability to use word study, decoding, and comprehension strategies; modeling effective strategies; integrating instruction across the language arts; grouping practices; use of a wide variety of genres; meeting the needs of both high-ability and struggling readers; and the ability to motivate students to value reading. The 33-item survey was reviewed for content validity by a panel of four experts in the field of reading and literacy instruction. A field test was then conducted to assess the clarity of wording of the items and instructions, appropriateness of the response scale and ease of administration. As a result, 11 questions were removed, leaving 22 final questions. Furthermore, the authors added instructions to the instrument that directed participants to "Please respond to each of the questions by considering the combination of your *current* ability, resources, and opportunity to do each of the following in your position" in an effort to heed Bandura's (1997) caution that self-efficacy instruments should capture assessments of current capability, rather than future potential or future intention, and to make the instrument context specific. The TSELI was shown to have high overall reliability rates ranging from Alpha .95 to .86.

Measures of teacher-efficacy interpretation problems.

One issue in the measurement of teacher efficacy is the degree of specificity of items in the scales and correspondence between tasks for self-efficacy measurement and outcome task.

Bandura (1997) emphasized the importance of content and context specificity in measuring perceptions of efficacy. Bandura (2001) cautioned that, "Scales of perceived self-efficacy must be tailored to the particular domains of functioning that are the object of interest" (p. 10). Through the years, a number of content-specific measures were created.

Table 2.3
Content-Specific Efficacy Measures

Subject	Authors
Math	Midgley et al., 1989
	Pajares & Miller, 1994
Special Education	Egyed & Short, 2006
	Meijer & Foster, 1988
Science	Riggs & Enoch, 1995
Writing	Graham, Harris, Fink, & MacArthur, 2010
Classroom Management	Emmer, 1990
Literacy	Tschannen-Moran & Johnson, 2011

Bandura (1997) developed a teacher efficacy scale with seven subscales to increase specificity in his studies: (a) efficacy to influence decision making, (b) efficacy to influence school resources, (c) efficacy to influence instructional efficacy, (d) efficacy to influence disciplinary efficacy, (e) efficacy to enlist parent involvement, (f) efficacy to create a positive school climate, and (g) efficacy to enlist community involvement. However, to date, this scale has not been published, and reliability and validity data are unavailable (Tschannen-Moran & Woolfolk Hoy, 2001).

Teacher self-assessment in terms of external standards of knowledge and skills has been a struggle associated with the measurement of teacher-efficacy. To date, few studies have examined this issue, but teachers' belief of the level of their content knowledge is not always especially accurate, as teachers tend to overestimate their knowledge and skills. Cunningham, Perry, Stanovich, and Stanovich (2004) asked primary teachers to assess their level of content knowledge of phonics and found that teachers tended to overestimate their knowledge and skills.

This overestimation of skills level could negatively affect self-efficacy beliefs, since personal knowledge and skills were not as strong as originally believed. Thus when measuring self-efficacy, the validity of teacher ratings must be considered.

Summary Methods of Measuring Self-Efficacy

Quantifying the construct of efficacy has required the development of measurement tools, which has proven to be a daunting task. The RAND organization (Armor et al., 1976) conducted the first studies to measure a teacher's perceptions of his or her own capabilities. The RAND researchers created and distributed a lengthy questionnaire to measure what they called teacher efficacy. The RAND study found a significant relationship between teacher efficacy and teachers' success in teaching reading to minority students in urban schools. Although the results of the RAND studies sparked interest in the construct of teacher efficacy, the reliability of the two-item scale concerned researchers. Consequently, attempts to develop longer, more comprehensive measures took place, including Guskey's (1981) Responsibility for Student Achievement, Rose and Medway's (1981) Teacher Locus of Control, the and Asthon vignettes (Ashton et al., 1984), none of which were widely accepted. One instrument that did gain acceptance, however, was the Gibson and Dembo Teacher Efficacy Scale (TES) (1984), which was used in almost half of the studies of teacher efficacy conducted up to 1998 (Ross, 1998).

Given that the measurement of efficacy was context specific, during the past several years, researchers developed a number of context specific tools. However, until recently, there still was a need for a literacy-specific measure. Tshannen-Moran and Johnson (2011) developed the Teacher Sense of Efficacy Scale. The authors based their development work on the Teacher Sense of Efficacy Scale (Tschannen-Moran & Woolfolk Hoy, 2001). The TSELI holds promise to further literacy-specific efficacy research.

Furthermore, interpreting the findings collected through various measures also requires attention. Teacher self-assessment in terms of external standards of knowledge and skills has been a struggle, since teachers tend to overestimate their own knowledge and skills. This overestimation may negatively affect self-efficacy beliefs if teachers' personal knowledge and skills were not in fact as strong as they believed. Thus the validity of teacher ratings must be considered.

Teacher Self-Efficacy and Literacy Instruction

While there has been an extensive review of teacher efficacy and its relation to various aspects of schooling, little research has attempted to uncover to what extent teacher efficacy predicted student achievement in the area of literacy. The studies that have been done, however, provided promising evidence that self-efficacy beliefs mattered in the realm of literacy instruction (Tschannen-Moran & Johnson, 2011). Armor et al. (1976) were the first to examine the impact of teacher efficacy on student achievement in reading. The authors' goal was to identify school and classroom policies and other factors that might increase the reading scores of inner-city children. Researchers in this study developed what has come to be known as the RAND measure of teacher efficacy, a two-item measure. The first item related to general teaching efficacy, or a teacher's belief about the impact of external factors compared to the influence of teachers and schools in general on student learning. The second item related to personal teaching efficacy, or a teacher's perceptions of how his or her teaching abilities influenced student learning. The Armor et al. study combined the two responses for an overall measure of teacher efficacy. Scores on the California Test of Basic Skills (CTBS) were used as measures of sixth- and seventh-graders' reading achievement in 32 district junior-high schools.

A strong correlation was found between teacher efficacy and teachers' success in teaching reading to minority students in an urban context.

Ashton and Webb (1986) studied 48 basic-skills teachers with students in grades 9-11. Scores in mathematics, language, and reading subtests of the Metropolitan Achievement Test (MAT) were obtained for the 1980 and 1981 school years for students who had been placed in special basic-skills classes for remediation and preparation for the test. Teacher efficacy was measured using a variety of instruments: a questionnaire that contained the two RAND items, an eight-item forced-choice measure of efficacy that has come to be known as the Webb efficacy measure, and a measure of personal teaching efficacy comprised of 15 vignettes now known as the Efficacy Vignettes. To measure how teacher efficacy, measured by the RAND and the Webb efficacy instruments, contributed to students' MAT reading-achievement scores, a stepwise multiple regression was computed controlling for the 1980 school-year scores. None of the correlations were found to be significant.

Another study examining the effect of teacher efficacy on students' reading achievement was conducted by Tranz and Gibson (1986), using Gibson and Dembo's Teacher Efficacy Scale. Fourteen elementary-school teachers and their students in grades 4-6 were studied. Class mean standard scores on CTBS reading, math, and language served as the measures of student achievement. In the Tranz and Gibson study, General Teaching Efficacy (GTE) and Personal Teaching Efficacy (PTE) were kept separate. In the analysis, these two factors had little correlation with each other ($r= .22$, $p=.22$) and correlated very differently across the three achievement areas. The only strong correlation involving reading achievement was with PTE ($r= .55$, $p<.05$). In stepwise multiple regression analysis with achievement scores as dependent measures, the variable of PTE accounted for a significant amount of variance in reading achievement ($p<.05$; $R^2=.30$).

The Gibson and Dembo instrument also was used by Anderson, Green, and Loewen (1988) to determine that PTE at the beginning of the school year predicted student achievement in language, reading, and math in grade 3, but not in grade 6. In the analysis, the teacher was assigned the class mean score for achievement. With the class as the unit of analysis, sample sizes were small, and the likelihood of finding statistical significance was reduced greatly. When a correlation was found between PTE and the achievement scores of third graders, two sets of stepwise multiple regressions were performed to determine which variables best accounted for the differences in student achievement. Teachers' sense of personal efficacy accounted for 54% of the variance for grade 3 students ($p<.05$; $R^2 = .54$).

In 1991, Borton expanded upon the research of Tranz and Gibson (1986) in a study of 79 third- and fourth-grade students and their teachers. Borton used the Gibson and Dembo Teacher Efficacy Scale for teacher efficacy and CTBS scores to measure student reading achievement. Neither teacher efficacy nor personal teaching efficacy emerged as significant predictor variables for student achievement scores on the reading portion of the CTBS, using a stepwise regression. Therefore, Borton's (1991) findings contradicted those of Gibson and Trancz (1986).

Hansen (2006) found that self-efficacy beliefs of English teachers differed across a range of English teaching tasks and competencies, highlighting that teachers purposefully may select to teach specific components of the English Language Arts (ELA) curriculum for which they feel they have greater efficacy, while neglecting areas in which they feel less confident.

Timperley and Phillips (2003) studied teachers who worked with disadvantaged students and initially found low expectations and an overall low level of teacher self-efficacy. After a six-month intervention in which teachers learned new and more powerful literacy teaching strategies and witnessed improved student outcomes as a result, teacher self-efficacy beliefs rose significantly. The type of training in reading instruction that pre-service teachers received was shown to have an effect on teaching in terms of differences in understandings, beliefs, and decision making. Beginning teachers who graduated from reading specialist and reading-embedded programs were found to be more willing to experiment with teaching methods than beginning teachers who graduated from general-education programs and tended to make decisions about teaching and learning in relation to external factors, such as available materials, mandates, and the wishes of administrators (Maloch, Flint, Eldridge, Harmon, Loven, & Fine et al., 2003). Self-efficacy beliefs also were found to influence the teaching of writing. Classroom practices in writing instruction of high self-efficacy teachers different greatly from that of low self-efficacy teachers (Graham, Harris, Fink, MacArthur, 2010). Teachers scoring the highest in self-efficacy reported spending significantly more time on writing each week and taught more writing processes and more grammar. Teachers with stronger self-efficacy beliefs also were more positive about the impact of natural learning methods.

Developing teacher self-efficacy for literacy instruction through Professional Development

The positive relationship between teacher self-efficacy and student achievement has led researchers to examine ways to improve teacher efficacy. To date, there is little empirical evidence regarding how to cultivate stronger self-efficacy for literacy instruction (Tshannen-Moran & Johnson, 2011). The teacher-efficacy studies that have been conducted show that efficacy can be developed through extended professional development, that engages teachers in meaningful activities that cause them to challenge their existing beliefs (Tschannen-Moran, &

McMaster, 2009). One such activity involves participation in teachers-as-readers groups and book clubs (Nathanson, Purslow, & Levitt, 2008) that met regularly to read and discuss the same literature, offering opportunities for teachers to explore their own literacy through sharing and reflecting. These teachers-as-readers groups gave teachers experience and confidence with book discussion, which could transfer to their own classroom and helped them learn from discussions with colleagues.

Cantrell and Callaway (2008) used a teacher-efficacy framework to describe the perceptions of high and low implementers of content literacy instruction using a year-long professional-development program. Interviews were used to collect teachers' comments and thoughts about their successes and challenges in affecting students' literacy development within the context of content-area instruction. The researchers found that teachers who embraced content literacy and implemented literacy techniques at high levels tended to exhibit higher general, personal, and collective teacher efficacy. High implementers believed that they had great influence on students' performance despite external difficulties and used multiple resources for finding strategies to meet students' literacy needs. In contrast, low implementers voiced that teachers had little power to make a real difference in students' literacy learning unless parents were involved and supportive. The low implementers were not as successful in working through the barriers they encountered as they attempted to implement content literacy strategies. Although low implementers initially did not seem to embrace the literacy strategies nor integrate them into their instruction, after professional development, they were more willing to consider the relevance of literacy integration and voiced a desire to learn more. This study suggested that teachers with lower efficacy might benefit from extended professional development to assist in shifting their beliefs. Guskey (1986) highlighted the importance of time in the process of shifts in teacher's beliefs.

Timperley and Phillips (2003) found generally low expectations and a low level of self-efficacy among teachers with students from disadvantaged backgrounds. After participating in a six-month intervention in which teachers learned new and more powerful literacy teaching strategies and observed improved student outcomes as a result, teacher self-efficacy beliefs increased significantly. The teachers' expectations of both themselves and their students increased regardless of student backgrounds.

Bandura (1997) highlighted four major influences on self-efficacy beliefs: vicarious experiences, verbal persuasion, physiological arousal, and mastery experiences. In an attempt to document the influence of the various sources of efficacy, researchers have conducted a number of studies. Yeung and Watkins (2000) found that prospective teachers' beliefs about their capability were influenced by the quality of supervision they received, verbal persuasion, and experiences during practice teaching, and mastery experiences. In a study of pre-service teachers, they found that teacher educator modeling positively influenced pre-service teachers' self-efficacy for effective literacy instruction (Johnson, 2010).

Tschannen-Moran and McMaster (2009) conducted a quasi-experimental study that tested the potency of different sources of efficacy beliefs. Participants implemented a new teaching strategy for beginning readers that used hand cues to help beginning readers bridge the gap between sounds and letter symbols. Three professional-development formats were studied: verbal persuasion (telling alone), the introduction of a limited mastery experience (practice with colleagues), and finally a full-mastery experience (coaching in the teachers' classrooms). The format that supported mastery experiences through follow-up coaching had the strongest effect on self-efficacy beliefs and the implementation of the new strategy. Furthermore, a significant number of teachers who participated in professional-development formats that included a demonstration with local students and a planning and practice session, but no follow-up coaching, experienced a decline in their self-efficacy belief. This study demonstrated that the development of self-efficacy beliefs was not a linear progression with new sources of self-efficacy adding to the gains in self-efficacy.

Researchers found the most powerful source of efficacy information for literacy teachers to be mastery experiences, or actual teaching experiences with students (Tschannen-Moran & MacFarlane, 2011). Self-efficacy beliefs were raised when a teacher witnessed improvement in student performances as a result of his or her teaching, which created optimism that future performances also will be successful. The increase in self-efficacy led to greater effort and persistence over time. Conversely, repeated failures lowered self-efficacy beliefs, leading to decreased motivation and resilience (Guskey, 1988; Ross, 19998; Tschannen-Moran et al., 1998). Fritz, Miler-Heyl, Kreutzer, and MacPhee (1995) found that after an in-service, teachers who attended the training had a higher self-efficacy than those in the control group even though both groups had equal self-efficacy pre-test scores.

Grade level also was found to affect efficacy results. In their 2011 study, Tschannen-Moran and Johnson found the elementary teachers demonstrated higher self-efficacy for literacy instruction than did middle-school teachers. These findings highlighted the need to provide middle-school teachers with greater preparation regarding literacy instruction.

Summary of Teacher Self-Efficacy and Literacy Instruction

Although there has been an extensive review of teacher efficacy and its relation to various aspects of schooling, little research exists regarding teacher efficacy and student literacy achievement. Studies demonstrated that self-efficacy beliefs affected literacy instruction (Tschannen-Moran & Johnson, 2011). Armor et al. (1976) were the first to examine the impact of teacher efficacy on student achievement in reading. A strong correlation between teacher efficacy and teachers' success in teaching reading to minority students in an urban context was found in the Armor et al. study. Subsequent studies provided mixed findings, but demonstrated that teacher efficacy did affect literacy instruction.

Recent research in the area of efficacy and literacy instruction indicated that appropriate professional-development opportunities could foster teacher efficacy for literacy instruction (Cantrell & Calloway, 2008; Fritz et al., 1995; Nathanson et al., 2008; Ross, 2007; Timperley & Phillips, 2003; Tschannen-Moran & McMaster, 2009; Yeung & Watkins, 2000). These professional-development studies have attempted to document the influence of Bandura's four areas affecting self-efficacy beliefs: vicarious experiences, verbal persuasion, physiological arousal, and mastery experiences. Mastery experiences proved to be the most powerful source of efficacy information for literacy teachers (Tschannen-Moran & MacFarlane, 2011).

CHAPTER THREE

Methodology

A description of the research design for this study, including the sample, data collection procedures, variables of interest, instruments used to measure the variables, and data analysis plan is provided in this chapter. Methodological points regarding the usage of various analytic techniques to address the respective study questions are outlined, specifically with regard to the implementation of analysis of variance (ANOVA) techniques and hierarchical linear modeling (HLM) as the primary analytic procedures utilized in this study.

Purpose of the Study

There is a large and growing body of literature on the effects of teachers' self-efficacy beliefs, and research has established a link between teachers' general perceptions of efficacy and student achievement. However, there is a dearth of literature regarding this effect specifically on reading achievement. While some previous research addressed the relationship between teacher efficacy and reading achievement, these studies did not employ subject-specific measures of efficacy. Bandura (2001) cautioned that, "Scales of perceived self-efficacy must be tailored to the particular domains of functioning that are the object of interest" (p. 1). Research on teachers' efficacy for literacy instruction has been hampered by the lack of a well-conceived, valid measure of the construct (Tschannen-Moran & Johnson, 2011). In addition to addressing the void in the literature focusing specifically on teachers' perceptions of self-efficacy for literacy instruction and student reading-achievement outcomes, this study also addressed variables that influence teachers' sense of efficacy for literacy instruction.

The purpose of this study was to examine the relationship between teacher and classroom variables on subject-specific efficacy and to investigate how teachers' efficacy beliefs for literacy instruction impacted reading achievement. It was hypothesized that high levels of efficacy for literacy instruction were associated with increased student-learning outcomes in reading in grades three through eight. Teacher self-efficacy for literacy instruction was measured by the Teachers' Sense of Efficacy for Literacy Instruction (TSELI) scale (Tschannen-Moran & Johnson, 2011). Reading achievement was measured by student scores on the Kansas Reading Assessment. In addition, contextual variables, including the number of years teaching, highest level of education, gender, and professional-development experiences of participating teachers, were collected in order to ascertain their influence on efficacy. Differences in classroom contextual variables based on student demographic information such as gender, free and reduced lunch status, special education status, ethnicity, and English language learner status were also evaluated with regard to their impact on teacher efficacy for literacy instruction.

Research Questions

The study focused on four research questions:

1. Does a teacher's level of sense of efficacy for literacy instruction decrease as grade levels increase?
2. How do teacher characteristics such as gender, the number of years teaching, education level, and professional development influence teachers' sense of efficacy for literacy instruction?
3. How do classroom-specific characteristics, such as gender, English language learner status, special-education status, free-and-reduced lunch status, and students' ethnicity influence teachers' sense of efficacy for literacy instruction?
4. How does a teacher's sense of efficacy for literacy instruction impact student reading achievement?

Research Design

Sample.

The sample for this study was drawn from three urban/suburban school districts in northeastern Kansas. Districts from metropolitan areas specifically were selected to mitigate potential differences that may have existed between larger, urban/suburban districts and smaller, rural districts. The school districts from which participants in the study sample originate are the second-, fourth-, and sixth-largest school districts in the state, respectively. Table 3.1 provides demographic information about the districts from which the sample was drawn.

Table 3.1
Demographic Information for Districts of Participants

District	Buildings			Enroll	Ethnicity			F/R Lunch	SPED	Grad Rate	AYP
	Elem	Mid	High		White	Black	Hisp				
A	34	5	5	27,876	68%	8%	15%	33%	10%	83%	No
B	20	9	6	21,435	80%	3%	4%	7%	9%	94%	Yes
C	18	6	4	14,194	50%	22%	16%	74%	15%	67%	No

Note: This information is based upon data from the 2010-11 academic year.

Teachers. Across the three districts, a total of 176 educators from grades three through eight completed the TSELI and the associated demographic questions. Table 3.2 details the number of teachers from each grade level who participated in the study.

Table 3.2
Surveys Collected from Grade-Level Teachers

Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Total
47	36	45	27	9	12	176

In addition to TSELI responses, teacher demographic data were also collected. Participating teachers were requested to respond to questions regarding their gender, the total number of years they had taught, and their highest level of education completed. The study sample was comprised of 21 (12%) male and 155 (88%) female teachers. The sample's median range of the total number of years teaching was between 11-15 years, while the mode was 6-10 years. One participant (1%) had earned a Doctorate, 62% had attained a Master's degree, and 37% had a bachelor's degree as their highest level of education completed. Table 3.3 details the teacher demographic information for all study participants by grade level.

Table 3.3
Teacher Demographic Information by Grade Level

Gr	Gender		Number of Years Teaching					Education Level				
	M	F	1-5	5-10	11-15	16-20	21+	B	B + 15	M	M + 15	D
3	1	46	10	13	10	6	8	4	15	12	16	0
4	6	30	6	13	7	3	7	6	14	4	12	0
5	7	38	8	15	11	3	8	6	13	13	12	1
6	6	21	8	5	4	7	3	0	8	10	9	0
7	1	8	2	1	0	2	4	1	1	3	4	0
8	0	12	2	4	1	3	2	1	2	4	5	0
Tot	21	155	36	51	33	24	32	18	53	46	58	1

Note: B= Bachelor's; M= Master's; D= Doctorate

The relationship between teacher demographic data and self-efficacy for literacy instruction was analyzed via mean comparisons (ANOVA) to assess whether teacher-level variables were associated with differences in levels of efficacy for literacy instruction. As a preliminary step, correlation coefficients were obtained examining the relationship between the ordinal demographic variables (total number of years teaching and education level of teachers) and the continuous TSELI scores. Pearson's *r* was utilized as the measure of association between these variables. Table 3.4 provides the obtained coefficients.

Table 3.4
Correlation Coefficients Between Teacher-Level Variables

	Education Level	TSELI Total
Years Teaching	.43**	.23**
Education Level		.16*

**Significant at the .01 level
*Significant at the .05 level

Both the number of years in the teaching profession and the education level of teachers were related to efficacy for literacy instruction scores. The direction of these moderately positive relationships was positive, and the nature of them was significant at the .01 and .05 levels, respectively. The correlation between the number of years teaching and education level was stronger in magnitude (r=.43, p<.01), indicating that educators who had taught longer tended to have higher education levels.

Basic descriptive statistics for TSELI scores for the study sample were calculated for each district as well by gender, by the total number of years teaching, and by the education level of participants for each grade level and the total sample. Table 3.5 displays the mean TSELI scores and average response value for all 22 TSELI items for all teachers within each district and for the total study sample.

Table 3.5
Mean TSELI Scores

	N	**Total Score**	**SD**	**Avg. Response**	**SD**
District A	65	148.4	20.4	6.8	.93
District B	24	168.1	18.4	7.6	.84
District C	87	144.4	21.8	6.6	.99
Total	176	149.1	22.2	6.8	1.0

From Table 3.5, average TSELI scores for District B ($\bar{X}$=168.1, SD=18.4) were considerably higher than the average reported efficacy for literacy instruction scores from Districts A ($\bar{X}$=148.4, SD=20.4) and C ($\bar{X}$=144.4, SD=21.8). The TSELI contains 22 items that ask teachers to respond on a scale ranging from 1 to 9. The average range of TSELI scores across districts was between 6.6 and 7.6, indicating that teachers typically perceived their level of efficacy for literacy instruction to be high.

Mean TSELI scores are presented in Table 3.6 by gender for each grade level and the total sample. Given the disproportionate sample sizes for the two gender groups, it was not prudent to draw any inferences on the basis of these descriptive statistics.

Table 3.6
Mean TSELI Scores by Gender

Grade	**Gender**			
	Male		Female	
	Mean (*n*)	SD	Mean	SD
3	164.0 (1)	NA	151.4 (46)	21.2
4	147.3 (6)	13.9	150.4 (30)	23.2
5	139.6 (7)	31.0	147.2 (38)	23.3
6	156.7 (6)	23.0	145.0 (21)	20.2
7	160.0 (1)	NA	151.8 (8)	24.7
8	NA	NA	149.5 (12)	23.7
Total	148.8	23.1	149.2	22.1

Note: Male *n*=21; Female *n*=155

When examining the average TSELI scores at each grade and for the total sample by the number of years teaching (Table 3.7), it appeared that efficacy for literacy instruction increased during the first two decades of a teacher's career before beginning to decline slightly. These differences in TSELI scores on the basis of length of tenure teaching were analyzed, and the results presented in Chapter 4.

Table 3.7
Mean TSELI Scores by Total Number of Years Teaching

Grade	**Total Number of Years Teaching**									
	1-5		6-10		11-15		16-20		21+	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
3	144.4	24.5	149.4	17.4	163.0	24.0	151.3	15.8	150.5	20.5
4	134.3	26.1	149.4	19.4	159.1	19.7	162.0	24.3	149.7	21.2
5	126.6	32.7	140.7	16.4	157.6	17.6	142.0	41.6	160.8	15.2
6	144.5	16.6	145.2	23.4	138.5	29.6	163.1	18.0	135.3	8.3
7	153.0	9.9	160.0	NA	NA	NA	141.0	38.2	156.5	28.3
8	136.0	11.3	144.5	32.5	138.0	NA	170.0	19.1	148.0	17.0
Total	138.8	24.4	146.3	19.1	156.7	21.9	156.4	23.1	152.1	19.4

Note: 1-5 *n*=36; 6-10 *n*=51; 11-15 *n*=33; 16-20 *n*=24; 21+ *n*=32

Table 3.8 presents mean TSELI scores at each grade level and for the total sample based on the education level of teacher participants. The categories included teachers who had completed a Bachelor's degree, those who had completed at least 15 hours of graduate coursework beyond a Bachelor's, those who had earned Master's degrees, those who had completed 15 hours of coursework beyond a Master's, and those who had completed a doctoral degree. For the total study sample, a visual inspection of the TSELI means for the entire study sample revealed that efficacy for literacy instruction scores tended to increase as education level increased.

Table 3.8
Mean TSELI Scores by Education Level

Grade	**Highest Education Level Attained**									
	B		B + 15		M		M +15		D	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
3	176.5	11.8	149.1	22.3	148.5	18.8	150.4	20.6	NA	NA
4	137.3	32.8	150.4	16.3	164.5	21.3	150.7	20.6	NA	NA
5	126.0	26.0	139.0	28.3	149.2	17.7	158.7	18.9	164.0	NA
6	NA	NA	134.8	14.1	156.8	14.4	148.7	27.4	NA	NA
7	160.0	NA	168.0	NA	147.3	12.1	151.0	34.9	NA	NA
8	144.0	NA	133.0	7.1	154.0	23.4	153.6	30.6	NA	NA
Total	143.9	30.5	144.6	21.6	152.3	17.6	152.2	22.65	164.0	NA

Note: Bachelor's *n*=18; Bachelor's +15= 53; Master's *n*=46; Master's + 15= 58; Doctoral *n*=1

Students. Students whose classroom teachers completed the TSELI were considered for inclusion in the study. All students of teachers who completed the TSELI and had a valid Kansas Assessment (2008) reading score subsequently were eligible to participate.

Kansas Assessment (2008) reading score performance data were supplied by the Assessment Director from each of the three participating districts in the study. Reading-achievement data were provided without any identifying information, such as student names or identification numbers. In addition to the reading-assessment scores, student demographics, including gender, ethnicity, English language learner status, special-education status, and free-or-reduced lunch status, were included. A class identification number that corresponded to the teacher-identification number also was included for each student. This variable was used to match teacher data from the TSELI with student data for analyses performed in the study.

A total of 3,814 students were eligible for inclusion and were retained for the study analyses. Table 3.9 details the characteristics of the study sample by each grade-level classroom that participated in the study.

Table 3.9
Characteristics of the Study Sample by Teacher

Teacher ID	District	Grade	TSELI SCORE	Reading score	N	Gender (%)		Ethnicity (%)		ELL (%)	SPED (%)	Lunch (%)
						M	F	Caucasian	Non-Caucasian			
1001	A	3	184	85.0	24	50	50	92	8	0	0	13
1002	A	3	126	77.5	26	50	50	50	50	23	4	73
1003	A	3	108	79.7	28	46	54	86	14	0	7	11
1004	A	3	150	83.1	18	61	39	72	28	0	0	17
1005	A	3	158	78.9	16	44	56	56	44	13	6	50
1006	A	3	162	86.2	21	48	52	76	24	0	5	10
1007	A	3	164	88.7	27	52	48	85	15	4	0	30
1008	A	3	148	86.2	26	54	46	65	35	8	4	23
1009	A	3	144	78.7	19	58	42	74	26	16	16	26
1010	A	3	146	88.7	25	56	44	84	16	0	8	28
1011	A	3	168	70.9	23	52	48	61	39	22	0	61
1012	A	3	136	79.4	28	50	50	50	50	11	4	46
1013	A	3	148	85.7	19	47	53	68	32	16	16	26
1014	A	3	122	81.0	15	47	53	53	47	27	7	53
1015	A	3	118	75.3	20	50	50	55	45	45	15	70
1016	A	3	158	80.8	19	42	58	68	32	11	11	63
1017	A	3	136	80.3	23	57	43	83	17	0	17	13
1018	A	3	178	77.8	21	57	43	62	38	0	19	48
1019	A	3	162	84.6	17	59	41	76	24	6	18	12
1020	B	3	160	84.6	25	56	44	60	40	12	8	16
1021	B	3	150	92.2	12	50	50	58	42	8	8	0
1022	B	3	186	88.1	24	50	50	75	25	8	8	8
1023	B	3	186	84.8	19	47	53	79	21	21	21	11
1024	B	3	158	86.2	27	59	41	93	7	0	7	0
1025	B	3	176	86.8	24	46	54	96	4	4	4	4
1026	C	3	140	75.2	24	58	42	50	50	0	13	79

1027	C	3	130	74.6	19	47	53	47	53	16	16	58
1028	C	3	144	70.6	18	61	39	22	78	6	11	94
1029	C	3	176	65.7	21	43	57	14	86	43	5	91
1030	C	3	132	76.9	14	57	43	50	50	0	21	71
1031	C	3	130	62.5	20	60	40	20	80	0	20	95
1032	C	3	178	71.4	24	54	46	42	58	58	4	92
1033	C	3	146	65.1	14	50	50	21	79	64	0	93
1034	C	3	144	82.7	22	50	50	32	68	5	9	73
1035	C	3	146	64.4	17	41	59	35	65	24	18	100
1036	C	3	192	74.9	17	41	59	41	59	6	24	35
1037	C	3	156	71.4	16	63	37	31	69	0	6	100
1038	C	3	180	78.5	21	52	48	62	38	0	14	52
1039	C	3	134	62.6	19	68	32	37	63	37	5	84
1040	C	3	154	76.6	22	41	59	36	64	5	9	73
1041	C	3	132	74.6	16	44	56	56	44	19	19	69
1042	A	4	140	83.7	23	57	43	65	35	13	0	9
1043	A	4	166	86.8	19	47	53	74	26	26	5	26
1044	A	4	134	81.7	20	55	45	45	55	25	15	85
1045	A	4	150	84.8	18	44	56	50	50	6	22	28
1046	A	4	174	83.2	24	33	67	67	33	8	17	33
1047	A	4	160	89.2	21	52	48	71	29	0	5	43
1048	A	4	146	82.4	15	40	60	67	33	20	13	47
1049	B	4	152	82.7	25	48	52	76	24	8	16	4
1050	B	4	140	88.4	27	52	48	89	11	0	11	0
1051	B	4	190	90.3	25	52	48	92	8	0	8	12
1052	B	4	146	89.3	25	56	44	72	28	4	4	0
1053	B	4	174	85.7	20	40	60	90	10	0	25	0
1054	B	4	192	87.1	22	55	45	82	18	0	9	5
1055	C	4	148	76.3	21	29	71	48	52	48	5	91
1056	C	4	146	76.0	21	52	48	38	62	24	14	100
1057	C	4	150	70.5	19	47	53	32	68	32	0	100

1058	C	4	156	67.4	18	50	50	28	72	0	28	94
1059	C	4	150	79.3	18	50	50	44	56	0	11	50
1060	C	4	122	76.1	18	44	56	28	72	0	0	78
1061	C	4	112	82.7	22	59	41	46	54	23	5	64
1062	C	4	142	83.7	21	52	48	52	48	0	0	57
1063	C	4	174	74.9	22	68	32	50	50	14	32	86
1064	C	4	178	81.0	19	63	37	53	47	0	21	84
1065	C	4	144	80.9	20	50	50	60	40	0	10	35
1066	C	4	138	77.5	21	43	57	14	86	0	10	91
1067	C	4	154	75.3	19	58	42	53	47	32	0	79
1068	C	4	138	77.8	14	57	43	43	57	0	7	64
1069	C	4	112	82.1	17	41	59	65	35	0	24	82
1070	C	4	172	69.4	17	53	47	24	76	6	6	100
1071	A	5	136	82.1	27	59	41	63	37	7	19	33
1072	A	5	154	81.8	17	47	53	35	65	24	12	82
1073	A	5	162	80.7	22	45	55	45	55	23	5	50
1074	A	5	154	80.1	23	57	43	57	43	13	9	57
1075	A	5	118	75.8	23	48	52	30	70	30	4	70
1076	A	5	184	74.5	28	46	54	39	61	25	11	54
1077	A	5	116	78.2	25	44	56	96	4	4	12	12
1078	A	5	120	82.3	26	58	42	46	54	4	15	54
1079	A	5	122	84.8	21	67	33	76	24	0	0	10
1080	A	5	144	83.9	23	35	65	70	30	0	4	35
1081	A	5	164	88.0	19	68	32	89	11	0	5	0
1082	A	5	126	88.2	26	50	50	96	4	4	8	31
1083	B	5	190	87.9	26	58	42	96	4	0	15	0
1084	B	5	192	91.6	21	57	43	95	5	0	14	0
1085	B	5	166	88.0	27	52	48	67	33	15	15	0
1086	B	5	156	90.0	21	52	48	81	19	5	19	5
1087	C	5	148	75.6	15	47	53	13	87	20	13	93
1088	C	5	160	81.3	22	41	59	77	23	0	0	73

1089	C	5	162	78.4	20	40	60	50	50	20	15	70
1090	C	5	152	72.2	18	50	50	50	50	0	0	83
1091	C	5	146	76.6	14	50	50	36	64	21	7	79
1092	C	5	156	65.9	20	55	45	25	75	0	15	100
1093	C	5	146	86.4	19	63	37	16	84	0	16	84
1094	A	6	142	87.8	24	38	63	92	8	4	0	4
1095	A	6	132	84.0	22	64	36	86	14	0	9	27
1096	A	6	158	82.9	20	50	50	90	10	0	5	15
1097	A	6	162	80.3	22	59	41	82	18	5	5	18
1098	A	6	166	80.8	21	38	62	48	52	10	14	67
1099	A	6	176	85.6	21	52	48	76	24	10	14	19
1100	A	6	190	87.4	23	48	52	87	13	9	13	13
1101	A	6	128	89.0	20	65	35	90	10	0	10	10
1102	A	6	142	81.9	24	50	50	83	17	0	13	21
1103	A	6	126	80.6	26	50	50	92	8	4	4	19
1104	A	6	138	88.1	26	46	54	92	8	0	4	8
1105	A	6	158	80.6	17	35	65	71	29	6	18	53
1106	A	6	156	87.2	19	47	53	68	32	0	5	26
1107	A	6	138	80.9	22	59	41	64	36	9	27	50
1108	A	6	186	80.1	22	45	55	91	9	0	23	36
1109	A	6	138	79.5	20	45	55	60	40	15	10	45
1110	A	6	114	81.0	23	48	52	52	48	4	9	43
1111	A	6	150	81.5	21	57	43	71	29	14	5	24
1112	B	6	142	84.1	53	53	47	87	13	2	2	9
1113	B	6	172	89.0	79	51	49	94	6	0	8	5
1114	B	6	172	87.5	43	63	37	91	9	0	0	9
1115	C	6	116	72.3	21	43	57	62	38	0	10	57
1116	C	6	154	73.9	59	49	51	56	44	0	22	56
1117	C	6	110	72.2	30	50	50	63	37	0	10	60
1118	C	6	140	75.1	62	50	50	39	61	2	13	89
1119	A	7	168	68.7	11	55	45	0	100	100	0	100

1120	A	7	136	86.7	46	52	48	70	30	11	0	30
1121	A	7	148	61.1	75	59	41	36	64	31	39	57
1122	A	7	146	86.0	44	55	45	75	25	2	11	11
1123	B	7	198	84.8	29	52	48	86	14	3	14	14
1124	B	7	150	89.0	87	47	53	76	24	1	9	1
1125	C	7	160	69.2	64	47	53	25	75	3	19	91
1126	C	7	114	80.9	62	47	53	48	52	0	0	55
1127	C	7	160	73.3	58	36	64	41	59	0	12	88
1128	A	8	136	60.1	34	47	53	35	65	56	44	56
1129	A	8	186	64.0	39	56	44	41	59	15	49	54
1130	A	8	110	74.9	93	52	48	44	56	10	13	55
1131	A	8	180	79.3	50	66	34	84	16	4	16	24
1132	A	8	152	58.4	59	56	44	31	69	19	47	61
1133	B	8	160	85.1	53	55	45	91	9	4	11	4
1134	B	8	144	86.4	83	54	46	87	13	0	1	2
1135	B	8	182	89.0	82	43	57	89	11	0	7	5
1136	C	8	138	82.0	60	47	53	53	47	3	0	58
1137	C	8	130	78.4	62	52	48	74	26	3	16	50
1138	C	8	128	68.9	39	56	44	23	77	5	21	90

Data collection.

Two sources of data were utilized in this study. Teacher self-efficacy for literacy instruction, as measured by the TSELI ,and demographic data were collected from teachers within the three districts that participated in the study. Student achievement scores, as measured by the Kansas Reading Assessment, and demographic data were obtained for eligible students within each teacher participant's classroom. These data were provided directly by the Director of Assessment from each of the participating districts.

The Director of Assessment from each of the three districts initially was approached regarding each respective district's participation in the study. The invitation to participate involved an in-person meeting, during which the researcher described the study and provided each Assessment Director with additional information detailing the study's purpose and data requirements. In addition, the researcher gave the Assessment Director a copy of the TSELI demographic questions (Appendices B-D). District C required approval from the district's Research and Evaluation committee. Assessment Directors from Districts A and B completed the University of Kansas Human Subjects Tutorial.

After confirming their respective district's involvement in the study, the Assessment Directors contacted the principal in each building throughout the district containing at least one of the grade levels treated in this study (three through eight) requesting that all educators involved in Reading instruction complete the TSELI and demographic questions. The purpose and objectives of the study were described to each building principal, along with the assurance that confidentiality and anonymity of teacher responses would be maintained.

Districts A and B elected to complete the surveys electronically using a web-based survey tool. Each classroom teacher solicited for participation in the study received a link to the

TSELI and demographic questions. Participants from these districts completed the surveys at their convenience, and upon submission of the electronic survey form, their involvement in the study was completed. District C opted to administer a paper-based version of the surveys. The researcher administered the paper-based form of the surveys to each eligible classroom teacher. Upon completion, teachers returned the surveys to the researcher, who manually entered the survey response data in a format mirroring that of data files containing the survey responses from Districts A and B.

Achievement and demographic data for each student of a teacher participating in the study was provided by the Assessment Director in each district. A data file structure that included all student variables requested was created by the researcher and provided to each Assessment Director. For districts A and B, whose teacher participants completed the surveys electronically, achievement and demographic data for eligible students was obtained on the basis of a class identification number. This identification variable was provided by teachers as part of the administered surveys and is a field included in the districts' student database. Using this variable maintained anonymity and allowed the researcher to pull data only for those students in the classes of teachers who participated in the study. For District C, whose respondents completed a paper-based version of the surveys, the Assessment Director was responsible for manually obtaining data for eligible students in the classes of participating teachers via the same class identification number variable.

Once all teacher and student data elements from each of the three participating districts were obtained, several data-management tasks were executed. The class identification variable was reviewed to ensure that each participating teacher was associated with a class of students contained in the student data files. Upon confirmation, all class identification numbers were transformed randomly to ensure anonymity of the data. In addition, all data were reviewed to confirm that students in the data file met all eligibility criteria.

For the TSELI and demographic information that was administered electronically (Districts A and B), all requested fields were completed by teachers as the computer-based format did not allow materials to be submitted unless all data elements had been entered. District C teachers who elected to participate in the study completed the same surveys, but their responses were collected via paper-based forms. Thirty-eight teachers submitted their completed forms with all elements completed except their teacher identification number. Unfortunately, there was no method for identifying these educators once their surveys were returned, which resulted in their exclusion from any analyses that required student-level data to be attached to their records.

All electronic files associated with and generated from the data and paper-based survey forms were stored in a secure location only accessible by the researcher. In addition, electronic files were contained on a password-protected computer only accessible by the researcher to maintain the confidentiality and integrity of the study data.

Study Variables of Interest

Student-level variables.

Student reading achievement. Student reading achievement was used as the dependent variable in the study. Reading achievement was measured by student scores from the Kansas Reading Assessment, administered as part of the Kansas Assessment Program (2008). All Kansas students in grades three through eight and grade 11 are required to participate in the Kansas Assessment Program.

Measurement instrument for student reading achievement. The Kansas Assessments are a program of the Kansas State Board of Education and are mandated by the Kansas State Legislature. In addition, the reading, mathematics, and science assessments are a part of the federal *No Child Left Behind* legislation. The assessment program is designed to (a) measure specific indicators within the Kansas Curricular Standards; (b) provide a building total score that measures Adequate Yearly Progress (AYP) and provides information for Quality Performance Accreditation (QPA); (c) report individual student scores along with the student's performance level; and (d) provide subscale and total scores that can be used in conjunction with local assessment scores to assist in improving a building or district's reading, mathematics, science, history/government, and writing programs (KSDE, 2008).

The Kansas Reading Assessment is a multiple-choice test that was developed to measure the targeted indicators of the Kansas Curricular Standards for the content area of reading. In each reading general-assessment test form, reading selections that represent different text types (narrative, expository, technical, or persuasive) are included based on those text types identified as appropriate in the grade-level test specifications. There are three parts to each grade-level reading assessment, with each part containing approximately 25-35 items. All students who are administered a Kansas Reading assessment receive a score and associated performance category. The metric for scores on the Kansas Reading assessment is based on the percentage of total items responded to correctly by an individual student, with scores ranging from 0 to 100. Scores from grade-level reading assessments are not linked to any other grade-level assessment above or below it. As a result, individual student progress on the Kansas Reading assessment cannot be monitored over time, and student scores cannot be compared across grades, as each grade-level assessment operates on different scales.

Technical information regarding the development, administration, and evaluation of the Kansas Reading assessments has been provided (Poggio, Yang, Irwin, Glasnapp, & Poggio, 2007). Score reliability estimates for the grade-level reading assessments have been calculated as Cronbach's alpha coefficients, with values ranging from a low of .88 to a high of .94 across all grade-level reading forms.

Students in grades three through eight and in grade 11 take the appropriate grade-level form of the Kansas Reading assessment during the state's mandated testing window, which occurs during the spring semester of the academic year (typically February to April). The Kansas Reading assessment data utilized in this study are from the 2010 and 2011 test administrations.

Student demographic variables. There is no shortage of research literature documenting that student learning is influenced by a host of factors. Many of these factors directly are related to what occurs in the classroom, but educational research consistently has found student achievement to be significantly influenced by contextual variables related to student demographic variables such as gender, ethnicity, and socioeconomic status. In a sentiment echoed by numerous researchers, Darling-Hammond (1997) contends that the contribution of school- and teacher-level variables to student achievement cannot be accurately measured without controlling for such contextual factors. For this reason, the student demographic variables of gender, ethnicity, English language learner status, special-education status, and free-and-reduced lunch status were included in this study. A brief description of each student demographic variable considered for control in this study follows.

Gender. This variable refers to the gender of a student as reported by the school district: male or female. In the entire study sample, there were 155 (88%) females and 21 (12%) males.

Ethnicity. This variable refers to a student's ethnic affiliation as reported by the school district: White, African-American, Hispanic, Native American, Asian, or Other. The ethnic makeup of the student sample in the study included 2367 (62%) Caucasian students, 449 (12%) African-American students, 603 (16%) Hispanic students, 50 (1%) Native American students, 136 (4%) Asian students, and 209 (6%) students classified as "Other." Overall, the study sample consisted of 62% Caucasian students and 38% non-Caucasian students.

English language learners. This variable refers to students whose first language is not English. In the study sample, 329 (9%) of students were classified as English language learners, while 3,485 (91%) spoke only English.

Special education. This variable refers to the special education status of students in the study sample. For a student to be classified as Special Education, he or she must have an Individual Education Plan (IEP). In the study sample, 448 (12%) of the students were classified as receiving special-education services, while 3,366 (88%) of students were not classified as such.

Free and reduced lunch. This variable served as a proxy for economic status of students. For the purpose of this study, this variable was divided into two categories: students who were eligible for the free-or-reduced lunch program and those who were not. In the student sample for this study, 1,618 (42%) students qualified for the free-or-reduced lunch program, and 2,196 (58%) did not qualify for the program.

Teacher-level variables.

Reading teachers' self-efficacy specific to literacy instruction was the independent variable for this study. Teacher efficacy was measured by scores based on responses to the

Teachers' Sense of Efficacy for Literacy Instruction (TSELI) scale (Tschannen-Moran & Johnson, 2011).

Measurement instrument for teacher efficacy. The TSELI (Appendix D) is an instrument that measures literacy-specific aspects of self-efficacy beliefs. The 22-item survey uses a Likert response scale on a nine-point continuum with 1 indicating "not at all," 3 indicating "very little," 5 indicating "some influence," 7 indicating "quite a bit," and 9 indicating "a great deal." Sample items include:

- To what extent can you help your students figure out unknown words when they are reading?
- To what extent can you provide specific, targeted feedback to students during oral reading?

Given that the TSELI is a relatively new measure of efficacy, the sections that follow detail the development of the scale and its psychometric properties of the instrument.

The need for a literacy-specific teacher efficacy measure.

Although a number of scales measure teacher efficacy, the lack of a literacy-specific measure has hindered the study of teacher efficacy relating to literacy. Bandura (1997) emphasized the importance of content and context specificity in measuring perceptions of efficacy. To address this gap in the measurement of efficacy for literacy instruction, Tschannen-Moran and Johnson (2011) developed the TSELI.

The development of the TSELI began with 33 items that were written specific to various aspects of literacy instruction and drew on the work of the NCTE and IRA *Standards for the English Language Arts* (1996) and the IRA *Standards for Reading Professionals* (2004). These items addressed aspects of literacy instruction such as the ability to use word study, decoding, and comprehension strategies, modeling effective strategies, integrating instruction across the language arts, grouping practices, use of a wide variety of genres, meeting the needs of both high ability and struggling readers, and the ability to motivate students to value reading. The 33-item survey was reviewed for content validity by a panel of four experts in the field of reading and literacy instruction. A field test was then conducted to assess the clarity of the items' wording and instructions, appropriateness of the response scale, and ease of administration. Furthermore, the authors added instructions to the instrument that directed participants to, "Please respond to each of the questions by considering the combination of your *current* ability, resources, and opportunity to do each of the following in your position" in an effort to keep with Bandura's (1997) caution that self-efficacy instruments should capture assessments of current capability rather than future potential or future intention, and to make the instrument context specific (Tschannen-Moran & Johnson, 2011).

Exploratory factor analyses were employed to pare the original pool of 33 items to those that demonstrated a coherent factor structure and to test the construct validity of the new scale. Eleven items were removed from the survey due to low communalities and factor coefficients. The factor structure of the remaining 22 items was analyzed, with all items loading on a single factor that explained 55% of the variance in teacher self-efficacy for literacy instruction scores that provided sufficient evidence of construct validity. The TSELI had a Cronbach's alpha reliability of .96.

The relationship between the subject-specific TSELI and the *Teacher Sense of Efficacy Scale* (TSES) (Tschannen-Moran & Woolfolk Hoy, 2001), an instrument that measures teachers' efficacy beliefs for more general aspects of teaching, was examined to determine whether the TSELI indeed measures more context- and subject-specific aspects of efficacy. Results of these analyses indicated that while the more general TSES and the more specific TSELI are related, they are distinct constructs that function differently.

Teacher demographic variables. Demographic variables focusing primarily on teacher experience were collected and considered in this study to evaluate their impact on teacher efficacy for literacy instruction. Previous research has examined the role of teacher variables and their relationship to self-efficacy beliefs. Bandura (1997) and others (Tschannen-Moran & Woolfolk Hoy, 2007) generally have found demographic factors to be weak predictors of teacher efficacy, but this research has focused on general aspects of the construct as opposed to the more subject- and context-specific factors measured by the TSELI. Findings such as these highlight the need for research using content-specific measures such as the TSELI. Demographic data surrounding the number of years teaching, education level, and gender of teachers were collected. A brief discussion of these variables follows.

Number of years teaching. This variable refers to the number of years a teacher taught: 1-5 years, 6-10 years, 11-15 years, 16-20 years, or 21 or more years. While previous research (Bandura, 1997; Tschannen-Moran & Woolfolk Hoy, 2007) has found that teachers who begin their careers with strong self-efficacy beliefs tend to maintain those beliefs throughout their careers, and those who begin with weak efficacy beliefs tend to persist in those beliefs, these studies focused on more general measures of teacher efficacy as opposed to specific aspects of the construct. In the teacher study sample, 36 (21%) teachers had taught for 1-5 years, 51 (29%) for 6-10 years, 33 (19%) for 11-15 years, 24 (14%) for 16-20 years, and 32 (18%) for 21 years or more.

Gender. This variable refers to whether the teacher is male or female. In previous teacher efficacy research, gender has been found to make a significant contribution to the formation of self-efficacy beliefs. However, it should be noted, that the amount of variance explained by the variable has been small enough to render this factor practically insignificant. The degree to which gender plays a role in more subject-specific measures of self-efficacy is not known, however. The teacher sample in this study consisted of 155 (88%) females and 21 (12%) males.

Education level. This variable refers to the education level of teachers in the study sample: Bachelor's degree, 15 hours of graduate coursework beyond a bachelor's degree, Master's degree, 15 hours of coursework beyond a Master's degree, or a Doctoral degree. In the study sample, 18 (10%) teachers had a Bachelor's degree, 53 (30%) had a Bachelor's plus 15 hours, 46 (26%) had a Master's degree, 58 (33%) had a Master's plus 15 hours, and one (1%) had a Doctoral degree.

Data Analysis

This study investigated the degree to which individual characteristics and contextual variables influenced a teacher's self-efficacy beliefs for literacy instruction, differences in teacher efficacy beliefs across grade levels, and the degree to which subject-specific teacher efficacy influenced student reading achievement. To assess whether there were significant differences in teachers' self-efficacy for literacy instruction given varying demographic, experience, and grade-level variables, several parametric analytic procedures were applied based on analysis of variance (ANOVA) techniques. Specific methods used to address whether individual characteristics influence teachers' self-efficacy beliefs for literacy instruction are described in this section.

Addressing the degree to which teacher efficacy influenced student reading achievement involved applying hierarchical, or multilevel, linear modeling (HLM). HLM is based on principles comparable to linear regression, whereby the predictive relationship between a set of independent variables and a dependent variable is evaluated. Because student achievement, as operationalized in this study, occurs at the student level, and teacher efficacy influences students at the classroom level, conventional methods of assessing this relationship become problematic because of the nested nature of data in this setting. A detailed discussion of the rationale, justification, and requirements for employing hierarchical modeling to address the impact of teacher efficacy for literacy instruction on student reading achievement follows.

Research Question 1: Does a teacher's levels of sense of efficacy for literacy instruction decrease as grade levels increase?

Research Question 2: What is the influence of teacher characteristics such as gender, the number of years teaching, education level, and professional development on teacher sense of efficacy for literacy instruction?

Research Question 3: What is the influence of classroom-specific characteristics such as gender, English language learner status, special education status, lunch status and the ethnic makeup of students on teacher sense of efficacy for literacy instruction?

Research questions one, two, and three addressed whether certain teacher-level and student-level variables influenced levels of efficacy for literacy instruction. To assess the degree to which demographic, contextual, and experience variables impacted the level of teachers' self-efficacy beliefs for literacy instruction, a series of ANOVA-based analyses were performed. For these analyses, the dependent variable of interest was teachers' efficacy for literacy instruction score as measured by the TSELI. Independent variables that were considered are detailed in Table 3.10.

Table 3.10
Independent Variables for Analysis of Research Questions 1-3

Variable	Values	Variable Type
Teacher Level	Research Questions 1-2	
Gender	Male; Female	Dichotomous
Teacher grade level	3; 4; 5; 6; 7; 8	Multiple levels
Years Teaching	1-5; 6-10; 11-15; 16-20; 21 or more	Multiple levels
Education level	Bachelor's; B+15; Master's; M+15; PhD	Multiple levels
Student Level	Research Question 3	
Gender	Male; Female	Dichotomous
Ethnicity	Caucasian; non-Caucasian	Dichotomous
ELL status	ELL; not-ELL	Dichotomous
SPED status	SPED; not-SPED	Dichotomous
Lunch status	Free/Reduced Lunch; Full pay	Dichotomous

To examine the impact of teacher characteristics (grade level, gender, number of years teaching, and education level) on efficacy for literacy instruction, analysis of variance (ANOVA) procedures were utilized. ANOVA was used to determine whether there were significant differences between the means of multiple independent (unrelated) groups. ANOVA was advantageous over performing multiple *t*-tests in the case of variables with more than two response categories, because it held Type I error rate in check (concluding that a significant difference existed when one, in fact, did not). ANOVA compares the means between the groups of interest and determined whether those means were significantly different from each other. Specifically, it tested the null hypothesis

$$H_0: \mu_1 = \mu_2 = \mu_3 \ldots \mu_k$$

where μ= the group mean and k= the number of groups. If the test returns a significant result, then the alternative hypothesis was accepted (H_A), which was that there were at least two group means that were significantly different from each other. ANOVA was an omnibus test statistic and did not reveal which specific groups were significantly different from each other, only that at least two groups were statistically different from one another. To determine which specific groups differed from each other, Tukey's honesty significant difference (HSD) post-hoc procedure was utilized. The Tukey test was not as conservative as other post-hoc tests, which meant it was more likely to detect differences if they actually existed.

For research questions one through three, one-way ANOVAs were conducted to assess the influence of each category of the independent variables on efficacy for literacy instruction. When significant mean differences between groups were observed for multiple variables of interest, two-way ANOVAs were executed. The two-way ANOVA procedure compared the mean differences between groups that had been split on two independent categorical variables (called factors). The two-way ANOVA not only assessed the difference in means for a factor collapsed over the levels of all other factors in the design (main effects), but also evaluated the presence of an interaction effect that occurred when the effect on one factor was not the same at the levels of another.

Effect sizes were obtained for all ANOVAs to indicate how much an independent variable affected the dependent variable in a study. In other words, effect size looked at how much variance in the dependent variable was a result of the independent variable. For ANOVA, the effect size calculated was η^2 (Eta squared), which is the ratio of between-groups variance to the total variance.

Research Question 4: What is the impact of teachers' sense of efficacy for literacy instruction on student reading achievement?

This question examined the impact of teacher efficacy for literacy instruction on student reading achievement. Teachers with high efficacy for literacy instruction were hypothesized to positively impact students' reading achievement as measured by scores on the Kansas Reading assessment. Classroom contextual variables based on student demographics including gender, ethnicity, English language learner status, special-education status, and participation in the free-or-reduced lunch program were controlled for in this analysis. Given the hierarchical nature of the data collected for this study, hierarchical linear modeling (HLM) was used to address this research question.

HLM analyses provided common, ANOVA-based statistical tests for model coefficients, main effects, and interaction terms to evaluate the relative contribution of variables to the model, including *t*-tests and chi-square tests. In addition, a series of model-fit statistics were provided so that a particular model could be compared against every other to determine which model fit the data most appropriately.

Characteristics of hierarchical linear modeling.

Data structures are often hierarchical in that variables frequently describe characteristics of individuals, who also are grouped into larger units consisting of a number of individuals. Furthermore, there are variables that describe characteristics of these higher-order units. In education, this hierarchical structure occurs naturally with students grouped into classes, classes grouped into schools, and schools grouped into districts.

Given the frequency and importance of hierarchical data in education settings, techniques are commonly applied that take hierarchical data into account. The first is to disaggregate all higher-order variables to the individual level. In this way, school, class, and teacher characteristics are assigned to each individual student, and subsequent analyses are performed at the individual level. The problem with this approach is that if all students in a class are assigned the same values on various variables, then the classical statistical assumption of independence of observations is violated. HLM, on the other hand, takes into account the interdependence of individual measures collected within the same organizational unit (the classroom, in this case) and assesses student-level, classroom-level, and between-level effects simultaneously.

The second commonly applied approach to treating hierarchical data structures is to aggregate the student-level variables to the higher-level variables and perform analyses on the higher level. In this manner, student characteristics are aggregated over classes, and analyses are performed at the class level. The main problem with this approach is that all within-group variation is disregarded when aggregating student data over classes, which can account for as much as 80-90% of the total variation (Raudenbush & Bryk, 2002). Thus important information is wasted when aggregating student-level data and analyzing at the higher-order level.

With both aggregating and disaggregating data in the approaches described, inappropriate inferences are often made to other levels of data. This so-called aggregation bias is one of the most commonly committed errors in classical statistical analysis (Raudenbush & Bryk, 2002). HLM, however, partitions the variance in a dependent variable into its within- and between-classroom components, making it possible to model the effect of the independent variable (teacher efficacy, in this analysis) on only the portion of the variance in the dependent variable (student reading achievement) occurring between classrooms while controlling for the influence of student demographic variables without aggregating these at the classroom level. In other words, it is possible to predict values of the dependent variable based on a function of predictor variables at more than one level.

An additional advantage of HLM is related to sample size. In classical statistical analysis, sample size must meet specific criteria. In actual classrooms, however, one class may have 18 students, while another may have 30 students. HLM allows analysis and interpretation of data from all classes, regardless of their size (Raudenbush & Bryk, 2002).

How teacher self-efficacy for literacy instruction impacted student reading achievement involves measuring variables at both the student and the teacher level. HLM is a multilevel modeling technique intended for nested data that is created in education settings. HLM is comparable to linear regression analysis, but considers the nested nature of educational data that would otherwise lead to violations of the basic assumptions of traditional linear analysis if these methods are applied to hierarchical data structures. Taking these factors into consideration, hierarchical linear modeling was adopted as the analytic procedure to address this study question.

The basic principle of HLM is comparable to that of Ordinary Least Squares (OLS) regression. At the student level (Level-1), an outcome variable is predicted as a function of a linear combination of one or more level-1 variables:

$$y_{ij} = \beta_{0j} + \beta_{1j}X_1 + \ldots + r_{ij}$$

where: β_{0j} represents the intercept of the group,

β_{1j} represents the slope of variable X_1 of group j, and

r_{ij} represents the residual for individual i within group j.

On subsequent levels, the Level-1 slope(s) and intercept become dependent variables being predicted from Level-2 variables, as shown in the following models:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}w_1 + \ldots + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}w_1 + \ldots + u_{1j}$$

where γ_{00} and γ_{10} are the intercepts and γ_{01} and γ_{01} are the slopes predicting β_{0j} and β_{1j}, respectively from variable w_1 . Through this process, the effects of Level-1 and Level-2 variables can be modeled to reflect the organizational structure of the data. By predicting the slopes as well as the intercepts (means), the differences in the relationship between Level-1and Level-2 can be explained.

Desired sample size for analysis.

An important methodological consideration for any statistical analysis is determining what sample size will achieve a desired level of power at a hypothesized effect size and an adopted level of significance. Power is the probability of rejecting the null hypothesis when a specific alternative hypothesis is true. In short, it is the probability of making a correct decision, and it is dependent on many factors, including the number of students within a class, the total number of classes analyzed, and the number of contextual variables being considered. Snijders and Bosker (1992) recommended that the number of classes analyzed be greater than or equal to 30 and that the number of students within a class be less than 28 in order to maintain the standard error below the desired level. Although others have offered their own recommendations with regard to adequate sample size, the literature on sample size in HLM is inconclusive (Snijders & Bosker, 1992; Raudenbush & Bryk, 2002).

HLM analysis design.

The focus of this research question was to examine the impact of teacher efficacy for literacy instruction on student reading achievement. The research question was investigated using hierarchical linear modeling with students nested within classrooms. The levels of the hierarchical models and the variables associated with each level are summarized in Table 3.11.

Table 3.11
Independent (Predictor) Variables for Reading Achievement (Dependent Variable)

Level	**Variable(s)**
Level-1 (student variables)	Gender Ethnicity English language learner status Special education status Free or reduced lunch status Grade
Level-2 (teacher variables)	Efficacy for literacy instruction (TSELI scores) District Education level Number of years teaching Gender

The logic of hierarchical linear modeling given the focus of this study involved building a series of models that incorporated Level-1 and Level-2 variables in an attempt to predict the dependent variable (reading achievement). The first step was to create an unconditional model, which included no Level-1 or Level-2 predictors. The unconditional model simply obtained the mean reading-achievement score for students across classes of teachers. This model indicated whether the class reading-achievement means varied significantly across classes, specifying the amount of variability at the student and teacher levels before any predictor or control variables were entered. The unconditional model served as a baseline to compare all subsequent models with predictor and control variables included to ascertain the contribution of each. The second step was to enter all teacher-level variables (Level-2) into the model in order to evaluate the relationship between each variable and reading achievement. Variables that contributed significant variance to student reading achievement were retained in subsequent models; those that did not contribute significant variance were dropped. The third family of models entailed entering the Level-1 variables (student-level variables) individually and investigating the contribution of all Level-1 variables in the model. Level-1 variables that were significant predictors of reading achievement were retained, and then subsequent model building included entering these significant predictors in various combinations to assess their predictive characteristics when joined together. The fourth series of models examined two-way interactions between all sets of Level-1 variables and Level-2 variables. The process of model building continued until all Level-1 variables and the sets of interactions between the total number of Level-2 variables retained had been evaluated. All models were judged according to various criteria, including individual significance and overall model fit. Table 3.12 provides a general overview of the hierarchical linear modeling steps that were implemented to address research question four.

Table 3.12
Summary of Structure for each Hierarchical Linear Model

Model Family (purpose)	Series	Variables Entered
1 (estimated the mean reading-achievement score across classes; specified the amount of variability at the student level and at the teacher level; provided a benchmark measure of deviance)	A	Null/unconditional model (no predictors)
	Level-2 Variables	
2 (assessed the contribution of each Level-2 variable and combinations thereof without including any Level-1 variables)	A	TSELI*
	B	District*
	C	Years Teaching (YT)
	D	Education Level (EL)
	E	TSELI, District
	F	TSELI, YT
	G	TSELI, EL
	H	District, YT
	I	District, EL
	J	YT, EL
	K	TSELI, District, YT
	L	TSELI, District, EL
	M	District, YT, EL
	N	TSELI, District, YT, EL
	Level-1 variables	
3 (estimated the effect for each Level-1 variable and the interaction of each Level-1 variable with each Level-2 variable)	A	Gender + retained Level-2 variables
	B	Grade + retained Level-2 variables
	C	Ethnicity + retained Level-2 variables*
	D	ELL status + retained Level-2 variables*
	E	SPED status + retained Level-2 variables*
	F	Lunch program status + retained Level-2 variables*
	G-	All combinations of Level-1 variables + retained Level-2 variables
	Interaction Terms	
4 (main effect and interaction models for Level-1 and Level-2 variables)	A-	Combinations of retained Level-1 variables with retained Level-2s
	Final Model	
5	A	TSELI, District, Ethnicity, ELL, SPED, F/R Lunch

Note: * refers to variables retained

Summary

This chapter outlined the purpose for the study as well as the research questions. Descriptions of the research design, including the sample, data collection, dependent and independent variables, and the instruments used to measure the variables were also provided. Additionally, a delineation of analyses conducted for each research question was provided.

CHAPTER FOUR

Results

This study investigated the relationship between various teacher and classroom characteristics and teachers' sense of efficacy for literacy instruction as well as the impact of efficacy for literacy instruction on student reading achievement. This chapter will present the study results in accordance with the methods detailed in the previous chapter, beginning with descriptive information for the study sample. Results for each of the four research questions are subsequently delineated. Data were analyzed using a combination of IBM SPSS 20.0 and the HLM 7 software package.

Purpose of Study

The purpose of this study was to investigate the effects of teacher efficacy in literacy instruction on third- through eighth-grade students' achievement on the reading test of the Kansas Assessment (2008). Higher levels of perceived teacher efficacy in literacy instruction were hypothesized to result in increases in the literacy achievement of students, controlling for student background variables such as grade, gender, ethnicity, and socioeconomic status. The influence of individual teacher characteristics such as number of years teaching, education level, and gender, as well as school contextual variables and the effect on student literacy achievement will be examined. Additionally, the effects of professional and personal development experiences on individual teacher efficacy were evaluated.

Research Questions

1. Does a teacher's level of sense of efficacy for literacy instruction decrease as grade levels increase?

2. How do teacher characteristics such as gender, the number of years teaching, education level, and professional development influence teachers' sense of efficacy for literacy instruction?
3. How do classroom-specific characteristics, such as gender, English language learner status, special-education status, free-and-reduced lunch status, and students' ethnicity influence teachers' sense of efficacy for literacy instruction?
4. How does a teacher's sense of efficacy for literacy instruction impact student reading achievement?

Descriptive Information

Sample.

Teachers. The overall teacher sample included 176 classroom teachers in grades 3-8 from the three districts selected for the study. Of the 76 teacher participants, 65 (37%) were from District A, 24 (14%) were from District B, and 87 (49%) were from District C. A majority of the participants were elementary (grades 3-6) classroom reading teachers (88%). Table 4.1 provides the number of teacher participants by grade level and district.

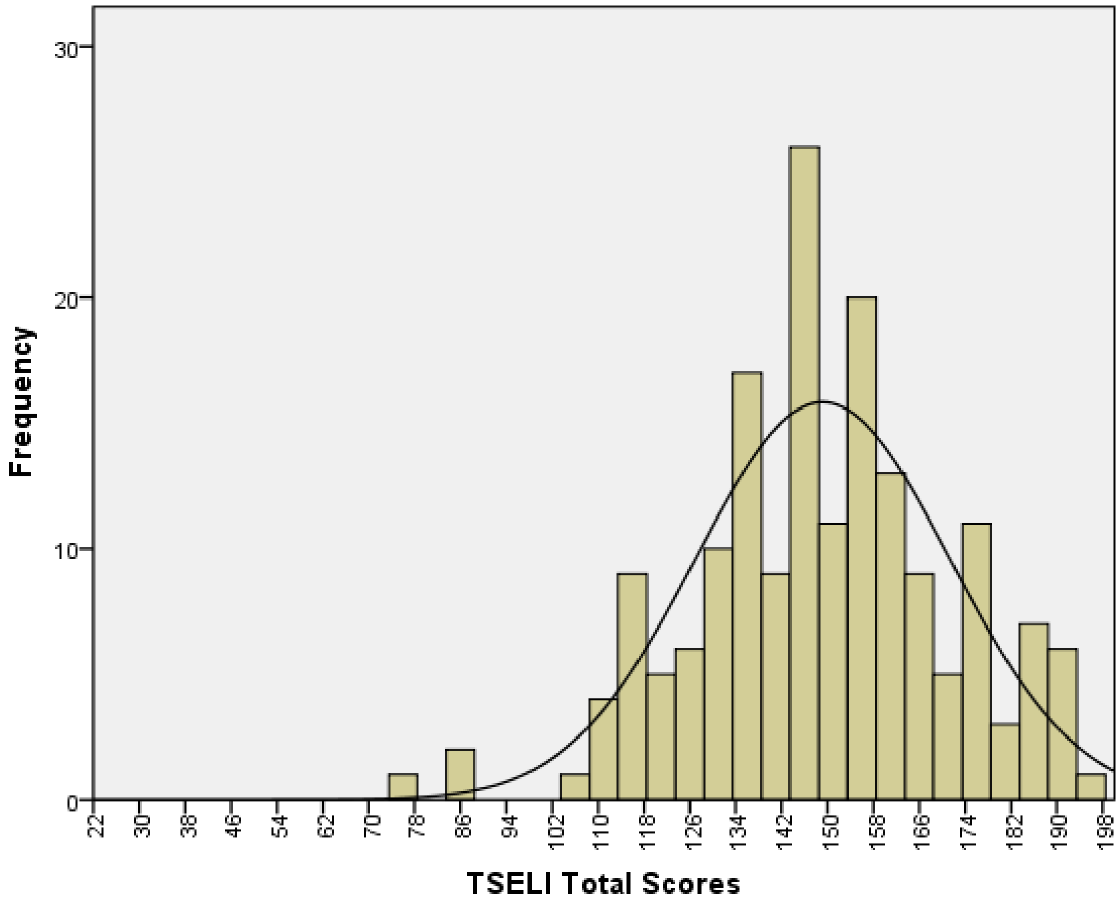
Table 4.1.
Teacher Participants by Grade Level and District

	District			
Grade	**A**	**B**	**C**	Total
3	19	6	22	47
4	7	6	23	36
5	12	4	29	45
6	18	3	6	27
7	3	2	4	9
8	6	3	3	12
Total	65	24	87	176

There were 155 (88%) female teachers in the sample and 21 (12%) males. Of the 176 teachers, 36 (21%) had taught between 1-5 years, 51 (29%) for 6-10 years, 33 (19%) for 11-15 years, 24 (14%) for 16-20 years, and 32 (18%) for more than 21 years. The education level of the teachers included 18 (10%) who had attained a Bachelor's as their highest degree attained, 53 (30%) who had at least 15 graduate hours beyond their Bachelor's degree, 46 (26%) who had a Master's degree, 58 (33%) who had at least 15 hours beyond a Master's, and one teacher (<1%) who had a doctorate.

To participate in this study, teachers were required to complete the Teacher Sense of Efficacy for Literacy Instruction (TSELI, Tschannen-Moran & Johnson, 2011). The TSELI consisted of 22 questions that were answered on a nine-point Likert scale. Scores on the TSELI ranged from a minimum of 22 to a maximum of 198. For all teachers in the sample, the average TSELI score was 149.1 (*SD*=22.2), with scores ranging from 76 to 198 (average response 3.5 and 9, respectively). The median TSELI total score for the sample was 148, while the most frequently occurring score was 146 (frequency=11, or 6%). The proximity of these three measures of central tendency to one another indicated that TSELI scores were approximately normally distributed within the study sample; however, the distribution was centered on the high end of the score scale (average response was 6.8, corresponding to the scale response category "Quite a bit"). Figure 4.1 details the distribution of TSELI total scores for the study sample.

Figure 4.1.
Distribution of TSELI Total Scores



The average TSELI scores were 148.4 (*SD*=20.4) for District A, 168.1 (*SD*=18.4) for District B, and 144.4 (*SD*=21.8) for District C. Tables 4.2 and 4.3 detail the average TSELI total score for each grade level across districts and for each grade level within districts, respectively.

Table 4.2.
TSELI Total Scores by Grade Across Districts

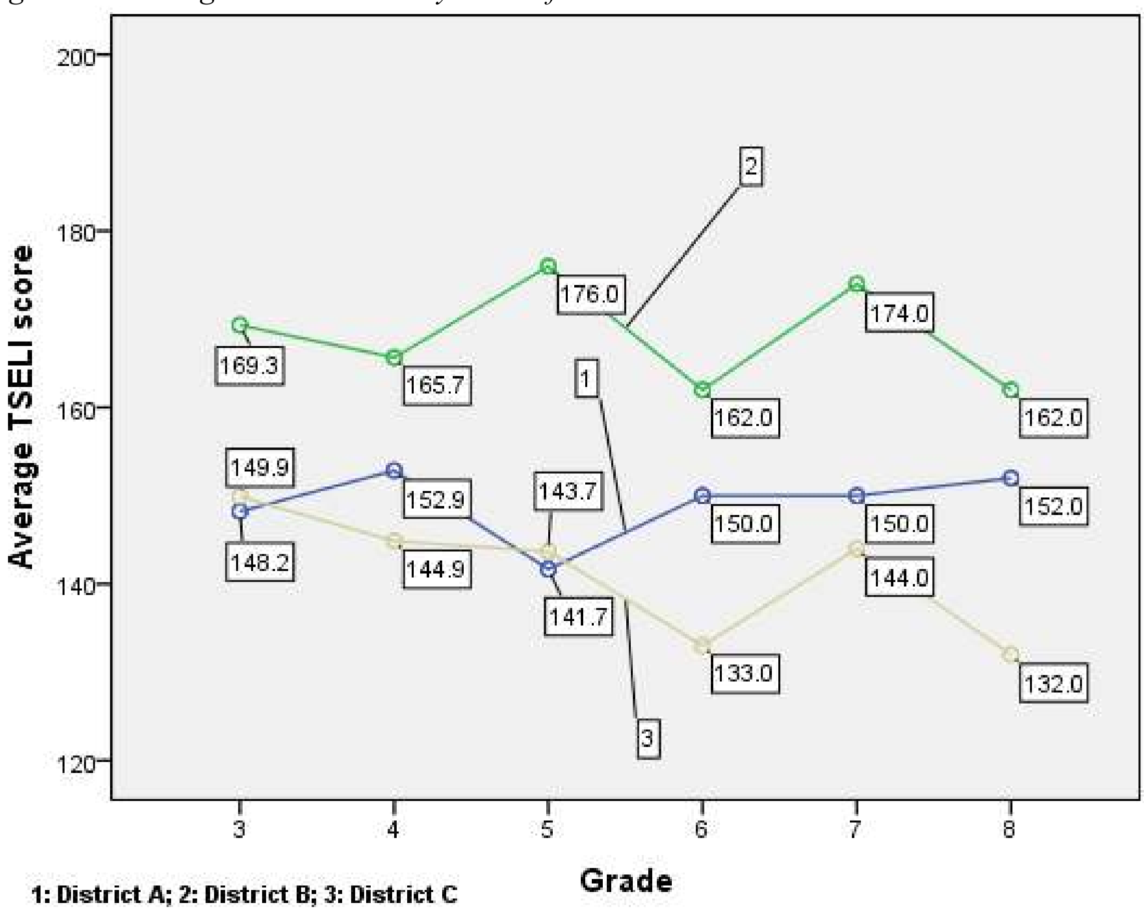
Grade	District	N	Mean	SD	Total
3	A	19	148.2	20.3	151.7
	B	6	169.3	15.4	
	C	22	149.9	21.4	
4	A	7	152.9	14.4	149.9
	B	6	165.7	22.7	
	C	24	144.9	22.0	
5	A	12	141.7	22.1	146.0
	B	4	176.0	17.8	
	C	29	143.7	23.8	
6	A	18	150.0	20.8	147.6
	B	3	162.0	17.3	
	C	6	133.0	16.9	
7	A	3	150.0	16.4	152.7
	B	2	174.0	33.9	
	C	4	144.0	21.7	
8	A	6	152.0	28.2	149.5
	B	3	162.0	19.1	
	C	3	132.0	5.3	

Table 4.3
TSELI Total Scores by Grade Within Districts

District	Grade	N	Mean	SD	Total
A	3	19	148.2	20.3	148.4
	4	7	152.9	14.4	
	5	12	141.7	22.1	
	6	18	150.0	20.8	
	7	3	150.0	16.4	
	8	6	152.0	28.2	
B	3	6	169.3	15.4	168.1
	4	6	165.7	22.7	
	5	4	176.0	17.8	
	6	3	162.0	17.3	
	7	2	174.0	33.9	
	8	3	162.0	19.1	
C	3	22	149.9	19.5	144.4
	4	23	144.9	21.6	
	5	29	143.7	46.1	
	6	6	133.0	16.5	
	7	4	144.0	21.7	
	8	3	132.0	5.3	

Table 4.2 indicates that between grades, variation in TSELI total scores appears low. That is, across grades looking at all districts, TSELI total scores are relatively stable (*range*=6.7 points, or 4% of the scale). However, within each grade there appears to be more variation in average TSELI total scores at the district level. Table 4.3 provides average TSELI scores for grade-level teachers within each district. While average TSELI grade-level scores within each district appear reasonably consistent, between districts there is a noticeable difference, as the average TSELI scores for District B teachers is markedly higher. Figure 4.2 illustrates the average grade-level TSELI scores for each district. A detailed analysis of these observed differences will be the focus of research questions one through three.

Figure 4.2.Average TSELI Scores by Grade for Each District



Students. Student data were supplied by the respective Assessment Directors from each of the three participating districts. These data were provided electronically according to a file structure pre-determined by the researcher without any identifying student information. Students were selected for inclusion in the study based on their reading teachers' responses on the TSELI. All students with a class identification number associated with a teacher responding to the TSELI were eligible for inclusion in the study. Once all TSELI and student data files were received from each of the participating districts, they were merged on the basis of matching teacher identification numbers to create the analysis sample.

After matching the student and teacher files, the researcher noted that student data for 38 teachers (58% of whom taught fifth grade) who responded to the TSELI from District C were not included in the file transmitted by that district's Director of Assessment. Student data were not transmitted for these 38 teachers in District C as these educators did not supply their unique teacher identification code on the TSELI survey materials. District C was the only participating district whose teachers did not complete the TSELI electronically, so teachers were able to return their survey materials without all fields completed. As there was no method for matching completed TSELI surveys to student data without the teacher identification code, it was determined that student data for these teachers could not be included in the study. The researcher determined, therefore, that analyses requiring only teacher data would be conducted using the complete set of teacher TSELI and demographic data (176 teachers) as detailed above. Research questions requiring the inclusion of student achievement and demographic data would be based on the complete set of teacher and student data (n=138 teachers).

A total of 3,814 students in grades three through eight were matched to 138 reading teachers who completed the TSELI. Of the 3,814 students, 1,866 (49%) were female, and 1,948

(51%) were male. In terms of ethnicity, 2,367 (62%) of the students were Caucasian and 1,447 (38%) were non-Caucasian. In the study sample, 329 (9%) of the students were classified as English Language Learners (ELL), 448 (12%) students were receiving special-education services (SPED), and 1,618 (42%) participated in the free/reduced lunch program. (Chapter 3, Table 3.8, details grade-level student-demographic information for the study sample by district.) Additionally, student reading achievement, as measured by scores on the Kansas Reading Assessment, was analyzed in this study. Average grade-level Kansas Reading Assessment scores are provided in Table 4.4.

Table 4.4.
Average Kansas Reading Assessment Scores for Study Sample

Grade	**District**	**N**	**District**		**Grade**		
			Mean	*SD*	N	M	*SD*
3	A	415	81.6	11.9	850	78.9	13.8
	B	131	86.7	8.2			
	C	304	71.9	15.1			
4	A	140	84.6	10.0	591	81.3	13.0
	B	144	87.3	9.6			
	C	307	77.0	14.0			
5	A	280	81.5	13.4	503	81.7	13.6
	B	95	89.2	8.0			
	C	128	76.7	14.7			
6	A	393	83.3	11.0	740	82.0	12.2
	B	175	87.1	8.3			
	C	172	73.8	14.2			
7	A	101	84.4	8.7	401	80.9	13.0
	B	116	88.0	7.8			
	C	184	74.4	14.5			
8	A	350	67.1	14.8	729	75.4	16.0
	B	218	87.1	8.6			
	C	161	77.5	15.7			

Delineation of Results

Research Question 1

Does a teacher's level of sense of efficacy for literacy instruction decrease as grade levels increase?

The one-way analysis of variance (ANOVA) procedure was used to determine whether there were significant differences between the means for each of the independent grade-level groups of teachers on the TSELI. For this analysis, the independent variable was the grade level of the teachers with five levels (grades 3 through 8) with the dependent variable being TSELI scores. The one-way ANOVA compared the TSELI means between the grade-level teacher groups and determined whether any of those means was significantly different from each other. Specifically for this research question, the ANOVA tested the null hypothesis:

$$H_0: \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7 = \mu_8$$

where μ_k was the mean TSELI score for the grade-level teachers in the study sample.

Testing of assumptions.

ANOVA makes the following assumptions regarding the probability distribution of the dependent variable:

1. The independent variable consists of two or more categorical independent groups
2. The dependent variable is either interval or ratio (continuous)
3. Independence of cases
4. The dependent variable is approximately normally distributed for each category of the independent variable
5. Equality of variances between the independent groups (homogeneity of variances)

The first three assumptions are study-design issues and were satisfied by the independent and dependent variables of interest in this analysis. The assumptions of normality and homogeneity of variances were assessed as a preliminary step prior to conducting this analysis. Normality is an underlying assumption for most parametric statistical tests and refers to the dependent variable being normally distributed in each group that is being compared. The degree to which TSELI scores were normally distributed for each of the grade-level teacher groups compared was addressed via numerical and graphical techniques. The Shapiro-Wilk Test was employed as the numerical means for assessing normality, as it is more appropriate for small sample sizes ($n < 50$) but also can handle sample sizes as large as 2000. Normal probability plots (Q-Q plots) also were used to graphically compare the distribution of TSELI scores at each grade level to the normal distribution. The straight line in the plots represents a perfectly normal distribution of TSELI scores for each grade-level group. The actual data for the study sample is represented by the points plotted along the line. If the data are normally distributed, then the data points will be close to the diagonal line. If the data points stray from the line in an obvious non-linear fashion, then the data are potentially distributed in a non-normal fashion. The various procedures employed to assess the statistical assumptions of normality and homogeneity of variances for the variables of interest in this analysis did not reveal any violations.

Results of analysis.

Table 4.5 provides the summary statistics from the ANOVA analysis, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for each separate grade-level group, as well as the total for all groups combined. Figure 4.3 provides a graphical representation of each grade-level TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.5
Summary Statistics: TSELI Scores by Grade Level

Grade	N	Mean	SD	Std. Error	95% CI Lower	95% CI Upper	Min	Max	Avg. Response
3	47	151.7	21.0	3.1	145.5	157.9	108	192	6.9
4	36	149.9	21.8	3.6	142.5	157.3	84	192	6.8
5	45	146.0	24.4	3.6	138.7	153.3	76	192	6.6
6	27	147.6	21.0	4.0	139.3	155.8	110	190	6.7
7	9	152.7	23.3	7.7	134.8	170.6	114	198	6.9
8	12	149.5	23.7	6.9	134.4	164.6	110	186	6.8
Total	176	149.1	22.2	1.7	145.8	152.4	76	198	6.8

Figure 4.3.
Average TSELI Score by Grade

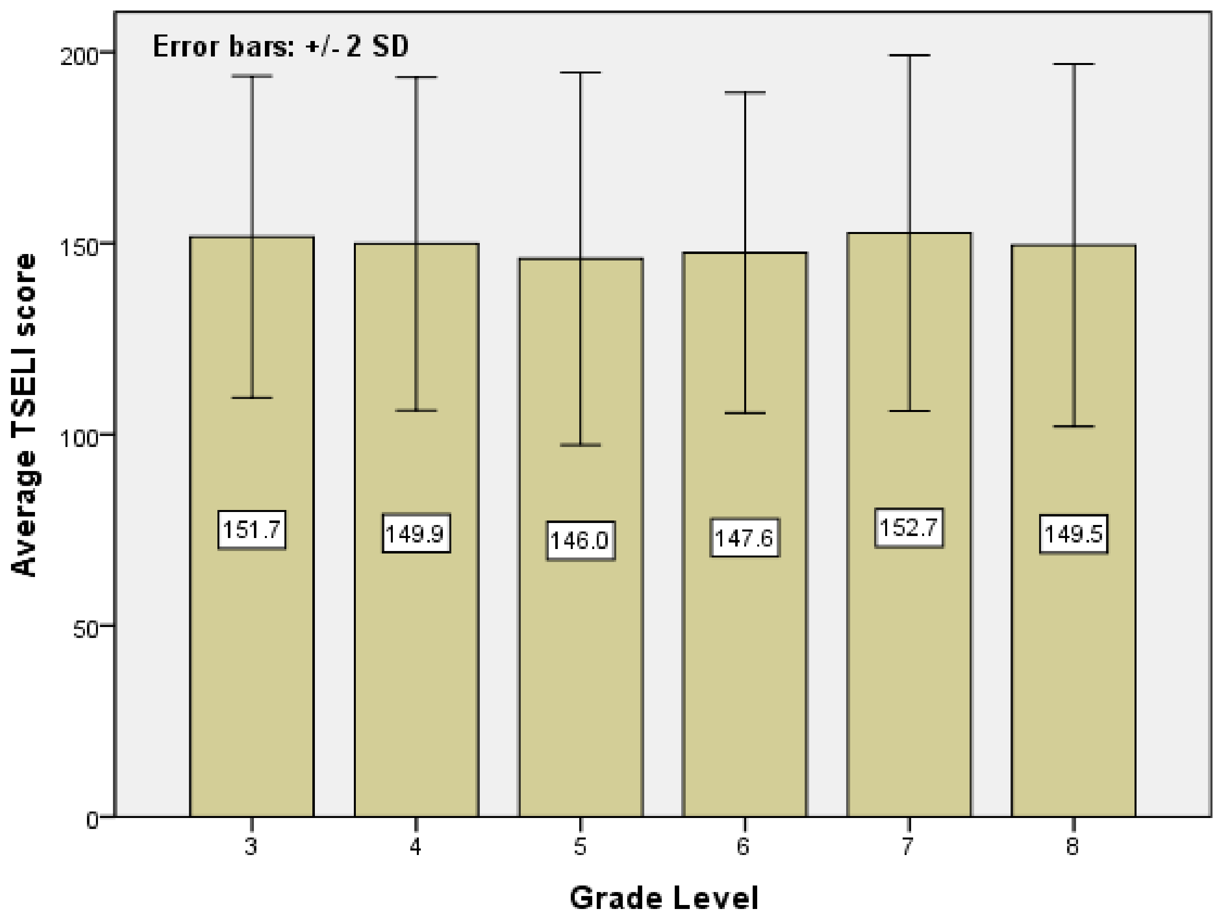


Table 4.6 displays the results from the ANOVA analysis. There was not a statistically significant difference between TSELI scores for grade-level groups as determined by the one-

way ANOVA ($F(5,170) = .38$, $p = .86$). The effect size ($\square^2$) of .01 was small according to Cohen's (1988) guidelines, indicating that only 1% of the variation in TSELI scores was due to differences in teachers' grade level.

Table 4.6.
One-way ANOVA: TSELI by Grade Level

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	953.68	5	190.74	.38	.86
Within Groups	85031.05	170	500.18		
Total	85984.73	175			

The one-way ANOVA was conducted to compare the effect of grade level on TSELI scores. Results from the one-way ANOVA indicated that teachers' sense of efficacy for literacy instruction did not significantly differ across grade levels when considering all teachers from the three participating districts in the study.

Follow-up analysis: Two-way ANOVA.

Figure 4.2 presented the grade-level TSELI means for teachers from each district in the study sample. Visual inspection of Figure 4.2 indicated considerable variation between districts in terms of grade level TSELI scores, so a two-way ANOVA was conducted to evaluate the mean differences in TSELI scores between groups of teachers split on two factors (independent variables): Grade and District. The two-way ANOVA was employed to determine if teachers' sense of efficacy for literacy instruction was influenced by their grade level and district in which they teach.

The same set of assumptions required of the data for a one-way ANOVA applied for a two-factor analysis of variance. Normality and heterogeneity of variance were assessed via numerical tests and graphical representations of the dependent variable. Tests of normality indicated that TSELI scores were distributed normally for all independent variables except for District B (grade 6) and District C (grade 5). These violations were not deemed to be serious, as the two-way ANOVA is considered a robust test against the normality assumption, meaning that it tolerates violations to this assumption with only a small effect on the Type I error rate. The heterogeneity of variance assumption was satisfied by these data.

Mean TSELI grade-level scores within each district are presented in Table 4.3. Results from the two-way ANOVA are shown in Table 4.7.

Table 4.7.
Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	13902.40	17	817.79	1.79	.03	.16
Intercept	2188947.08	1	2188947.08	4798.04	.00	.97
Grade	944.61	5	188.92	.41	.84	.01
District	9979.97	2	4989.98	10.94	.00	.12
Grade * District	2616.79	10	261.68	.57	.83	.04
Error	72082.33	158	456.22			
Total	4000516.00	176				
Corrected Total	85984.73	175				

The outlined rows in Table 4.7 inform whether there were significant mean differences between the groups for the independent variables Grade and District and for their interaction, Grade*District. The main effect of Grade on TSELI scores yielded the same result as the one-way ANOVA: there was no significant difference between grade-level groups on TSELI scores (F(5,158)=.41, p=.84). The main effect for District indicated that there was a significant difference between TSELI scores across participating districts (F(2,158)=10.94, p<.01), indicating that that mean TSELI score was greater for District B teachers ($\bar{X}$= 168.2) than for both District A ($\bar{X}$= 149.1, p<.01) and District C teachers ($\bar{X}$= 141.2, p<.01). The difference in mean TSELI scores between District A ($\bar{X}$= 149.1) and District C ($\bar{X}$= 141.2) was not found to be significant (p=.09). Table 4.8 presents the results from the Tukey honestly significant difference (HSD) *post-hoc* test that was employed to confirm where the differences in mean TSELI scores occurred between the study districts. It should be noted that the average values of the dependent variable in a two-way ANOVA are based on the estimated marginal means when the design is unbalanced (number of cases in each cell is not equal). Estimated marginal means adjust for the covariate (Grade) in a two-factor ANOVA by reporting the means of the dependent variable (TSELI score) for each level of the factor (District) at the mean value of the covariate (Grade). Overall, the District effect accounted for 12% of the variation in TSELI scores.

Table 4.8.
Pairwise Comparisons for District Effect

(I) District Teach	(J) District Teach	Mean Difference (I-J)	Std. Error	Sig.	95% CI for the Difference	
					Lower bound	Upper bound
District A	B	-19.0	5.7	.00	-30.4	-7.7
	C	7.9	4.6	.09	-1.3	17.1
District B	A	19.0	5.7	.00	7.7	30.4
	C	26.9	5.7	.00	15.5	38.3
District C	A	-7.9	4.6	.09	-17.1	1.3
	B	-26.9	5.7	.00	-38.3	-15.5

Finally, there was not a significant interaction between the effects of Grade and District on sense of efficacy for literacy instruction ($F(10,158) = .57$, $p=.83$), meaning that there was not a difference in grade-level TSELI scores depending on the district in which the study participants taught. Taken together, the results from the one-way and two-way ANOVAs did not show a significant difference in grade-level TSELI scores overall or within the respective districts.

Research question 2.

What is the influence of teacher characteristics, such as gender, the number of years teaching, and education level, and professional development, on teacher sense of efficacy for literacy instruction?

The effect of individual teacher variables on sense of efficacy for literacy instruction was assessed using ANOVA procedures to determine whether there were significant differences between the means of TSELI scores for various groups based on educator characteristics such as gender, the number of years teaching, and education level. Additionally, teacher perceptions of the quality of their undergraduate training, quality and frequency of district professional development, and involvement in personal enrichment activities (such as participation in a book club) were evaluated. For this series of analyses, the independent variables of interest included:

- Gender: Male or female
- Highest education level completed: Bachelor's degree, Bachelor's + 15 graduate credit hours, Master's degree, Master's + 15 hours, or Ph.D.
- Quality of undergraduate teacher-training program: Self-reported Likert scale ranging from 1 (low) to 10 (high)
- Quality of district professional-development sessions: Self-reported Likert scale ranging from 1 (low) to 10 (high)
- Frequency of professional-development sessions during academic year
- Participation in a book club: Yes or No

Influence of gender on efficacy for literacy instruction.

For this research question, a one-way ANOVA was implemented to compare the TSELI means between male and female teachers. It should be noted that a one-way ANOVA for an independent variable with only two levels is analogous to an independent samples *t*-test, and therefore produces identical results regardless of the method utilized. For this analysis, the ANOVA tested the null hypothesis :

$$H_0\colon \mu_M = \mu_F$$

where μ_k was the mean TSELI score for the two gender groups in the study sample.

Testing of assumptions.

The assumptions of ANOVA (detailed in section 4.4.1.1) were tested prior to executing this analysis. No violations of normality or homogeneous variances were observed.

Results of analysis.

Table 4.9 provides the summary statistics from the ANOVA analysis, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for both gender groups and for both groups combined. Figure 4.4 provides a graphical representation of each gender group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.9
Summary Statistics: TSELI Scores by Gender

Gender	**N**	**Mean**	**SD**	**Std. Error**	**95% CI**		**Min**	**Max**	**Avg. Response**
					Lower	**Upper**			
Female	155	149.2	22.1	1.8	145.7	152.7	76	198	6.8
Male	21	148.8	23.1	5.1	138.3	159.3	85	186	6.8
Total	176	149.1	22.2	1.7	145.8	152.4	76	198	6.8

Figure 4.4 Average TSELI Score for Male and Female Teachers

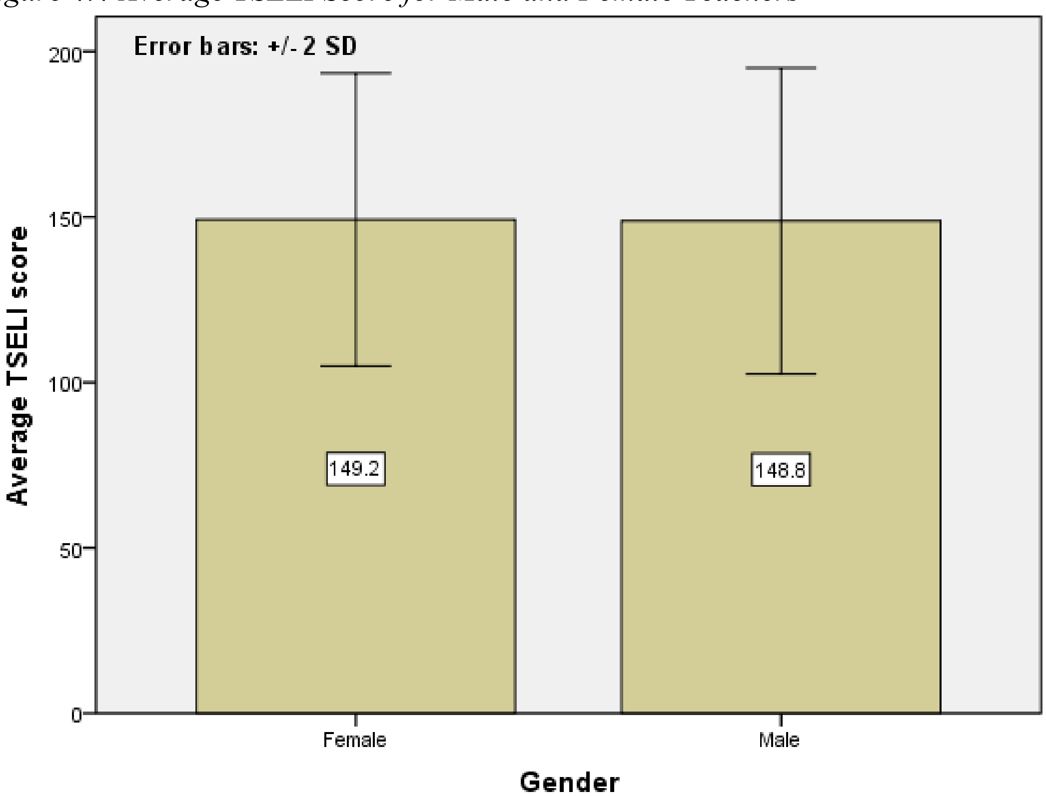


Table 4.10 displays the results from the ANOVA analysis. There was not a statistically significant difference between TSELI scores for males and females as determined by the one-

way ANOVA ($F(1,174) = .005$, $p = .94$). The effect size ($\square^2$) was 0, indicating that gender did not account for any variation in TSELI scores.

Table 4.10.
One-way ANOVA: TSELI Scores by Gender

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	2.55	1	2.55	.005	.94
Within Groups	85982.18	174	494.15		
Total	85984.73	175			

The one-way ANOVA was conducted to compare the effect of gender on TSELI scores. Results from the one-way ANOVA indicated that teachers' sense of efficacy for literacy instruction did not significantly differ for males and females.

Influence of the number of years teaching on efficacy for literacy instruction.

For this research question, the one-way analysis of variance procedure was used to determine whether differences existed between levels of efficacy for literacy instruction depending on the number of years teachers had been in the profession. Data were collected from teachers in the study sample regarding length of tenure as an educator: 1-5 years (group 1), 6-10 years (2), 11-15 years (3), 16-20 years (4), and 21 or more years (5). For this analysis, the one-way ANOVA tested the null hypothesis:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

where μ_k was the mean TSELI score corresponding to the groups of teachers in the study sample based on the number of years they had taught.

Testing of assumptions.

The assumptions of ANOVA (detailed in section 4.4.1.1) were tested prior to executing this analysis. No violations of normality or homogeneity of variances were observed.

Results of analysis.

Table 4.11 provides the summary statistics from the one-way ANOVA, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for all groups based on length of time teaching as well as for all groups combined. Figure 4.6 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.11
Summary statistics: TSELI Scores by Number of Years Teaching

Length Teaching	**N**	**Mean**	**SD**	**Std. Error**	**95% CI**		**Min**	**Max**	**Avg. Response**
					Lower	**Upper**			
1-5 yrs	36	138.8	24.4	4.1	130.6	147.1	76	188	6.3
6-10 yrs	51	146.3	19.1	2.7	140.9	151.6	110	186	6.7
11-15 yrs	33	156.7	21.9	3.8	149.0	164.5	110	192	7.1
16-20 yrs	24	156.4	23.1	4.7	146.7	166.2	114	190	7.1
21+ yrrs	32	152.1	19.4	3.4	145.1	159.1	108	198	6.9
Total	176	149.1	22.2	1.7	145.8	152.4	76	198	6.8

Figure 4.5.
Average TSELI Score for Groups Based on Length of Time Teaching

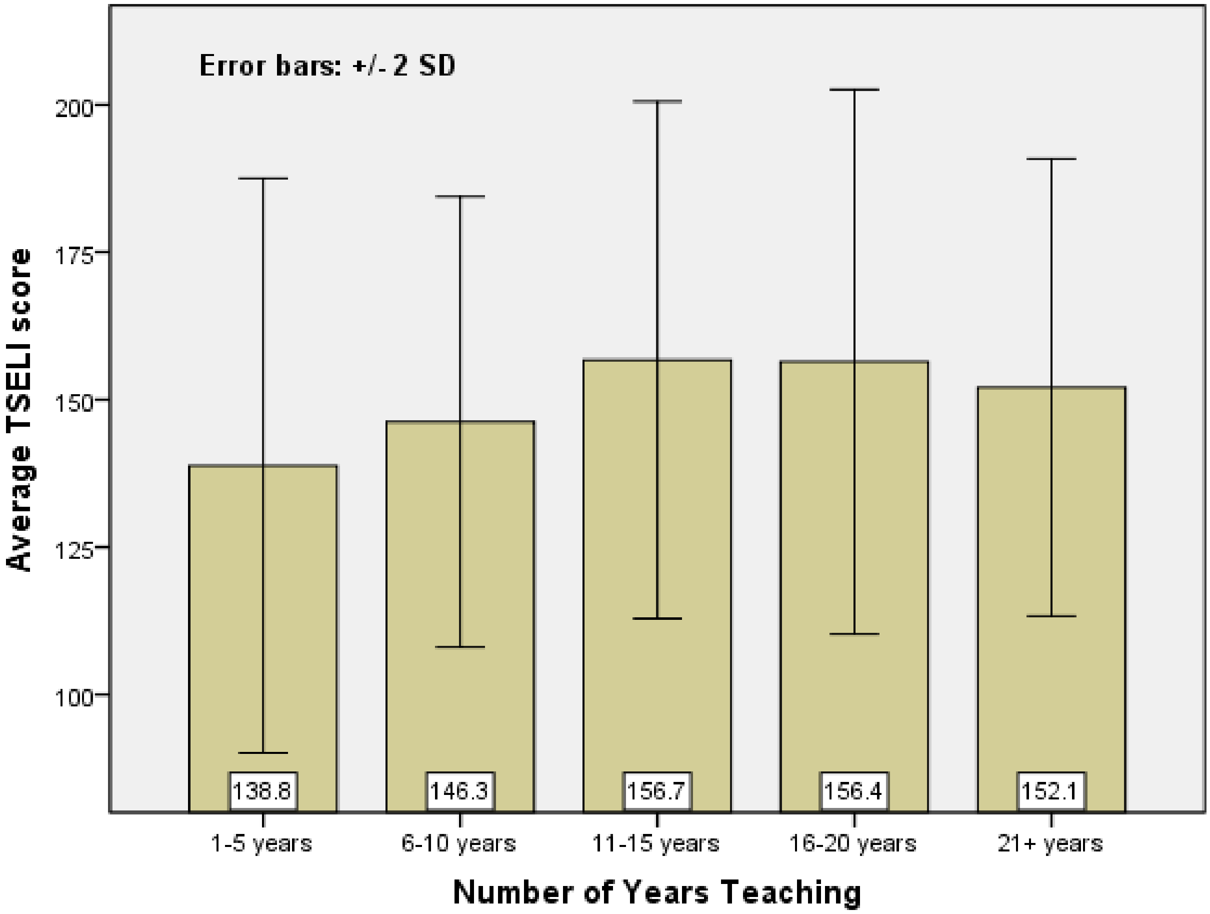


Table 4.12 displays the results from the ANOVA analysis. There was a statistically significant difference in TSELI scores between groups as determined by the one-way ANOVA ($F(4,171) = 4.21$, $p < .01$). The effect size (□^{2}) was .09, indicating that length of time teaching accounted for 9% of the variation in TSELI scores.

Table 4.12
One-way ANOVA: TSELI Scores by Number of Years Teaching

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7713.15	4	1928.29	4.21	.003
Within Groups	78271.58	171	457.73		
Total	85984.73	175			

The one-way ANOVA was conducted to compare the effect of length of time teaching on TSELI scores. Results from the analysis indicated that teachers' sense of efficacy for literacy instruction significantly differed based on how long they had taught. The ANOVA test is an omnibus test, informing whether an overall difference exists between the groups. However, ANOVA does not indicate which specific groups differed. Therefore, the Tukey HSD *post-hoc* test was used to confirm where differences occurred between groups. The Tukey test was selected because it is not as conservative (meaning that it is more likely to detect differences if they exist) as the other post-hoc tests that can be used when the data meet the assumption of homogeneity of variances. Table 4.13 shows the Tukey post-hoc test results for the different levels years of teaching experience.

Table 4.13.
Multiple Comparisons for Effect of Length of Time Teaching

(I) Years Teach	(J) Years Teach	Mean Difference (I-J)	Std. Error	Sig.	95% CI for the Difference	
					Lower bound	Upper bound
1-5 years	6-10 years	-7.45	4.66	.50	-20.29	5.39
	11-15 years	-17.92	5.16	.01*	-32.14	-3.71
	16-20 years	-17.61	5.64	.02*	-33.16	-2.07
	21+ years	-13.26	5.20	.05*	-27.59	1.07
6-10 years	1-5 years	7.45	4.66	.50	-5.39	20.29
	11-15 years	-10.47	4.78	.19	-23.65	2.71
	16-20 years	-10.16	5.30	.31	-24.76	4.44
	21+ years	-5.81	4.82	.75	-19.11	7.49
11-15 years	1-5 years	17.92	5.16	.01*	3.71	32.14
	6-10 years	10.47	4.78	.19	-2.71	23.65
	16-20 years	.31	5.74	1.00	-15.51	16.14
	21+ years	4.66	5.31	.90	-9.97	19.30
16-20 years	1-5 years	17.61	5.64	.02*	2.07	33.16
	6-10 years	10.16	5.30	.31	-4.44	24.76
	11-15 years	-.31	5.74	1.00	-16.14	15.51
	21+ years	4.35	5.78	.94	-11.57	20.28
21+ years	1-5 years	13.26	5.20	.05*	-1.07	27.59
	6-10 years	5.81	4.82	.75	-7.49	19.11
	11-15 years	-4.66	5.31	.90	-19.30	9.97
	16-20 years	-4.35	5.78	.94	-20.28	11.57

*The mean difference is significant at the 0.05 level

From the results of the Tukey post-hoc tests presented in Table 4.13, significant mean TSELI differences existed between educators who had taught between 1-5 years and those who had taught between 11-15 years (p<.05), 16-20 years (p<.05), and 21+ years (p<.05). There were no other statistically significant differences between other combinations of groups.

Influence of education level on efficacy for literacy instruction.

For this research question, a one-way ANOVA was conducted to determine whether statistically significant differences existed between levels of efficacy for literacy instruction and the education level of teachers. Data were collected from teachers in the study sample regarding their education level: Bachelor's degree (Group 1), Bachelor's + 15 graduate credit hours (Group 2), Master's degree (Group 3), Master's + 15 credit hours (Group 4), and Ph.D. degree (Group 5). For this analysis, the one-way ANOVA tested the null hypothesis:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

where μ_k was the mean TSELI score corresponding to the groups of teachers in the study sample based on education level.

Testing of assumptions.

The assumptions of ANOVA (detailed in section 4.4.1.1) were tested prior to executing this analysis. No violations of normality or homogeneity of variances were observed.

Results of analysis.

Table 4.14 provides the summary statistics from the one-way ANOVA, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for all groups based on education level and for all groups combined. Figure 4.7 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.14.
Summary statistics: TSELI Scores by Education Level

Education Level	N	Mean	SD	Std. Error	95% CI		Min	Max	Avg. Response
					Lower	Upper			
Bachelor's	18	143.9	30.5	7.2	128.7	159.1	84	188	6.5
Bachelor's +15	53	144.5	21.6	3.0	138.6	150.5	76	192	6.6
Master's	46	152.3	17.6	2.6	147.0	157.5	118	190	6.9
Master's + 15	58	152.2	22.6	3.0	146.3	158.1	108	198	6.9
Ph.D	1	164.0	.	.	.	.	164	164	7.5
Total	176	149.1	22.2	1.7	145.8	152.4	76	198	6.8

Figure 4.6
Average TSELI Score for Groups Based on Education Level

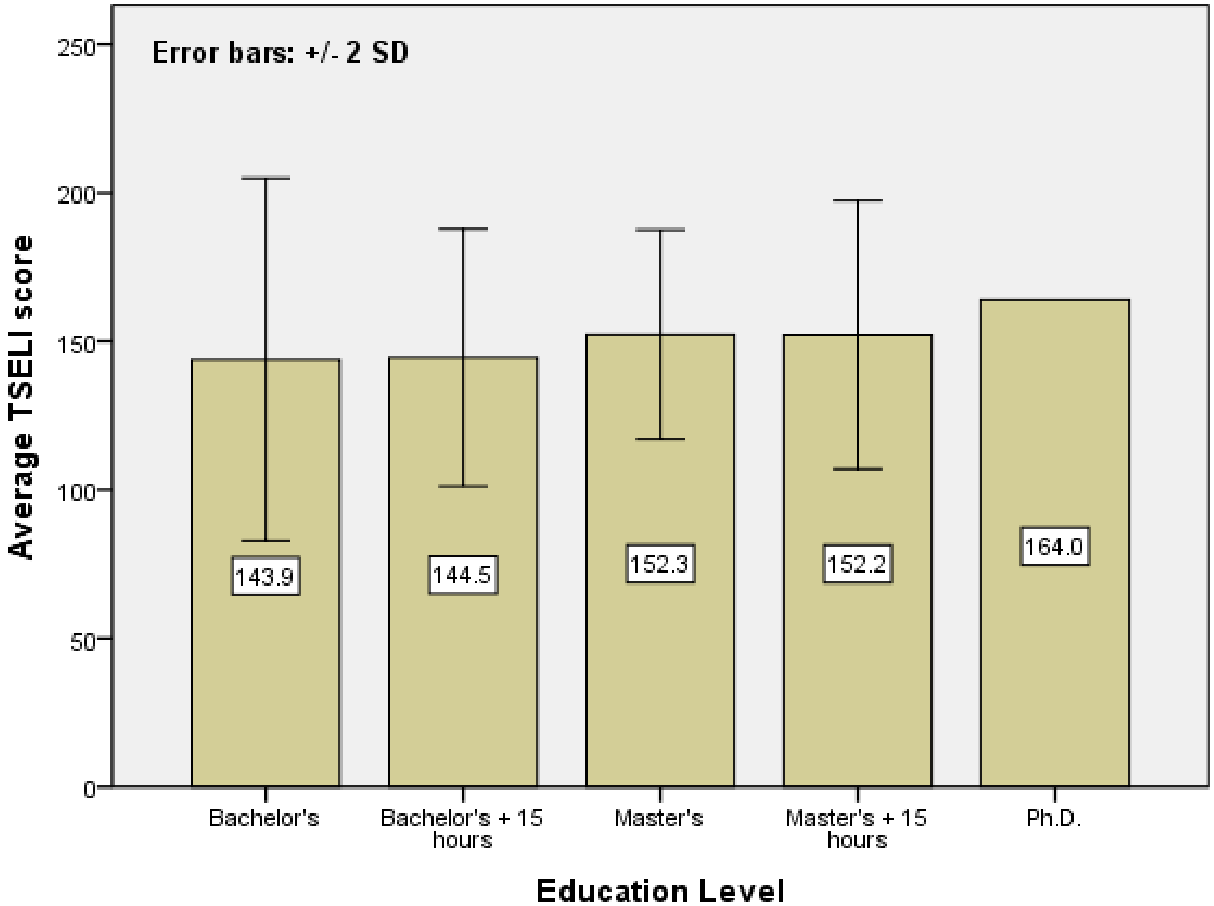


Table 4.15 displays the results from the one-way ANOVA. There was not a statistically significant difference in TSELI scores between groups as determined by the one-way ANOVA. ($F(4,171) = 1.46$, $p = .22$). The effect size (□^{2}) was .03, indicating that teachers' education levels accounted for 3% of the variation in TSELI scores.

Table 4.15.
One-way ANOVA: TSELI Scores by Education Level

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2834.97	4	708.74	1.46	.22
Within Groups	83149.75	171	486.26		
Total	85984.73	175			

The one-way ANOVA was conducted to compare the effect of education level on TSELI scores. Results from the analysis indicated that teachers' sense of efficacy for literacy instruction did not significantly differ based on education level. A subsequent analysis, however, was conducted whereby the education level variable was recoded to reflect the highest level of education completed by teachers in the study sample. Teachers were categorized as having either an undergraduate degree or graduate degree in the following manner:

- Group 1, undergraduate degree: Bachelor's and Bachelor's + 15 graduate hours
- Group 2, graduate degree: Master's, Master's + 15 graduate hours, and Ph.D.

Table 4.16 provides the summary statistics from the ANOVA (analogous to an independent samples *t*-test in this case), including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for the two groups and for all groups combined. Figure 4.8 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.16
Summary statistics: TSELI Scores by Degree

Education Level	N	Mean	SD	Std. Error	95% CI		Min	Max	Avg. Response
					Lower	Upper			
Undergraduate	71	144.4	23.9	2.8	138.7	150.1	76	192	6.6
Graduate	105	152.4	20.4	2.0	148.4	156.3	108	198	6.9
Total	176	149.1	22.2	1.7	145.8	152.4	76	198	6.8

Figure 4.7
Average TSELI Score for Groups Based on Highest Education Level Completed

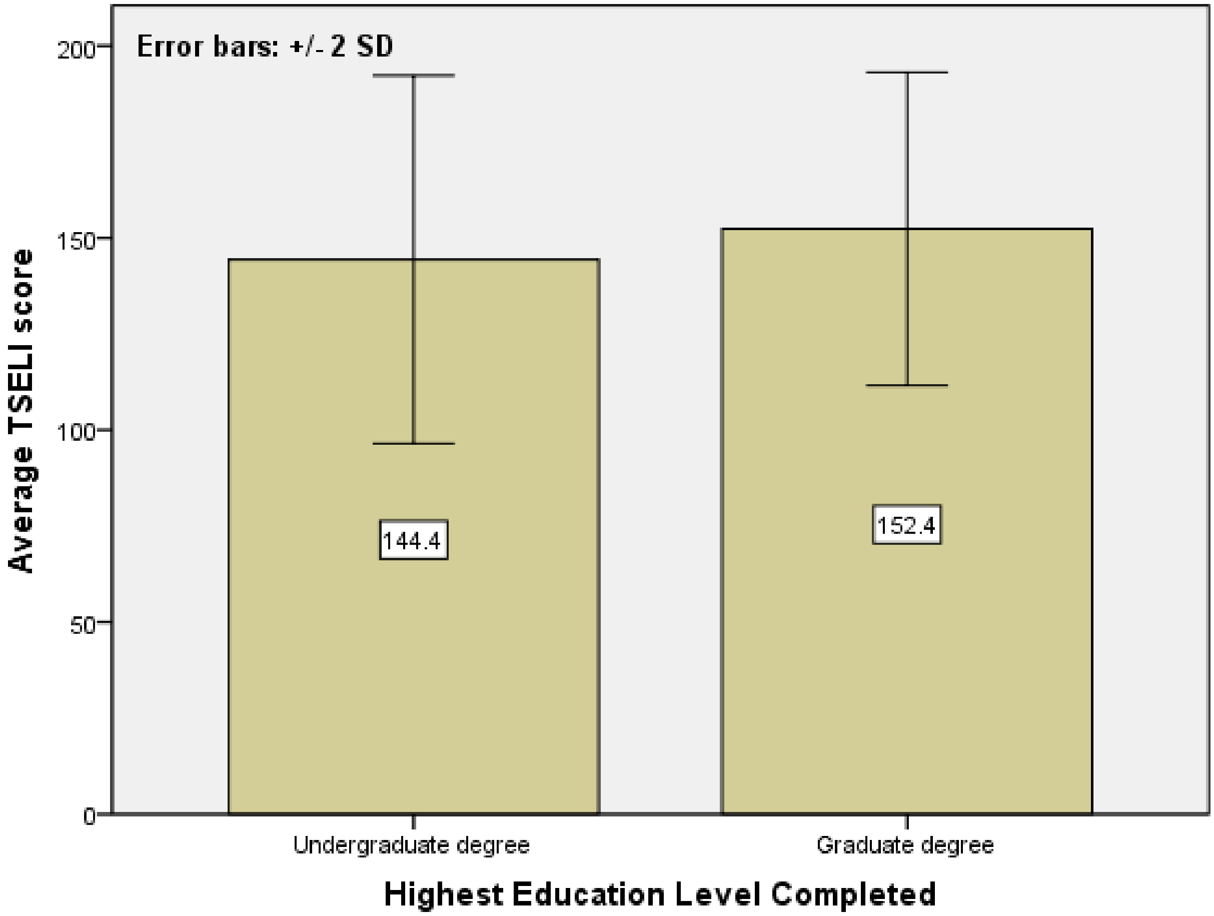


Table 4.17 displays the results from the one-way ANOVA. There was a statistically significant difference in TSELI scores between groups as determined by the one-way ANOVA ($F(1,174) = 5.62$, $p<.05$). The effect size (□^{2}) was .03, indicating that teachers' highest education level completed accounted for 3% of the variation in TSELI scores.

Table 4.17
One-way ANOVA table: TSELI Scores by Degree

	Sum of Squares	**Df**	**Mean Square**	**F**	**Sig.**
Between Groups	2692.03	1	2692.03	5.62	.02
Within Groups	83292.69	174	478.69		
Total	85984.73	175			

The teacher education-level data were recoded to reflect highest education level completed, and a one-way ANOVA was conducted to compare the effect of this variable on TSELI scores. Results from the analysis indicated that teachers' sense of efficacy for literacy instruction was significantly higher for teachers with graduate degrees.

Two-way ANOVA: Influence of years teaching and education level on efficacy

A two-way analysis of variance was conducted to evaluate whether an interaction effect was present among the two independent variables found to significantly impact efficacy for literacy instruction scores, number of years teaching and highest education level completed. An interaction exists between two factors if the effect of one factor depends on the levels of the second factor. Specifically, this analysis sought to address whether the effects of number of years teaching depend on the highest education level completed by teachers in the study sample.

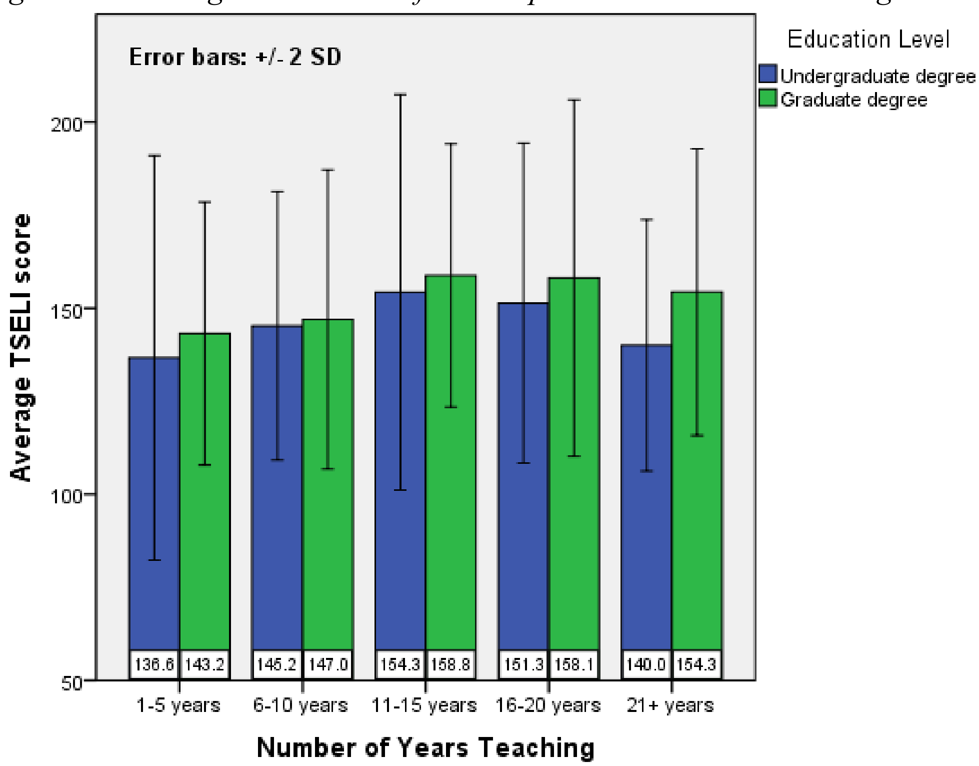
Results of analysis.

Table 4.18 provides the summary statistics from the two-way ANOVA, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for all groups based on education level as well as for all groups combined. Figure 4.8 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.18
Summary statistics: TSELI Scores by Years Teaching and Education Level

Years Teaching	Education Level	N	Mean	SD	Avg. response
1-5 years	Undergraduate	24	136.6	27.2	6.2
	Graduate	12	143.2	17.7	6.5
	Total	36	138.8	24.4	6.3
6-10 years	Undergraduate	21	145.2	18.0	6.6
	Graduate	30	147.0	20.1	6.7
	Total	51	146.3	19.1	6.7
11-15 years	Undergraduate	15	154.3	26.6	7.0
	Graduate	18	158.8	17.7	7.2
	Total	33	156.7	21.9	7.1
16-20 years	Undergraduate	6	151.3	21.5	6.9
	Graduate	18	158.1	23.9	7.2
	Total	24	156.4	23.1	7.1
21+ years	Undergraduate	5	140.0	16.9	6.4
	Graduate	27	154.3	19.3	7.0
	Total	32	152.1	19.4	6.9
Total	Undergraduate	71	144.4	23.9	6.6
	Graduate	105	152.4	20.4	6.9
	Total	176	149.1	22.2	6.8

Figure 4.8. Average TSELI Score for Groups Based on Years Teaching and Education Level



Results from the two-way ANOVA are shown in Table 4.19.

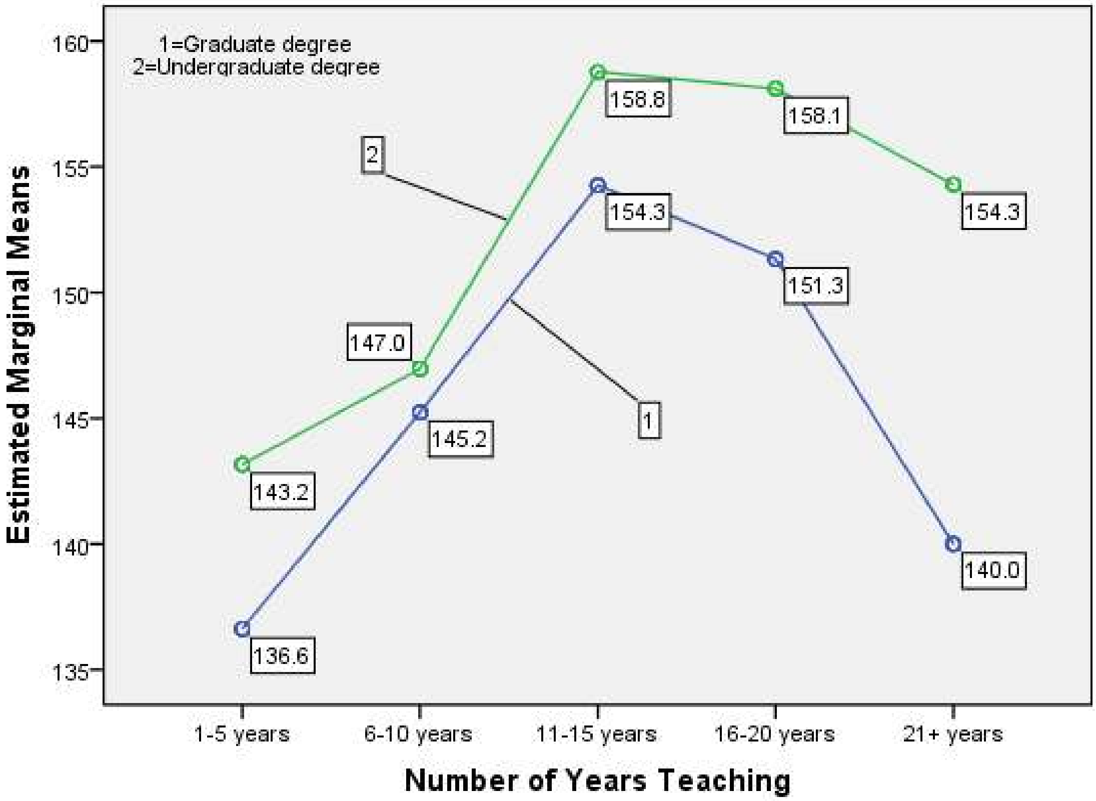
Table 4.19.
Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	9327.87	9	1036.43	2.24	.02	.11
Intercept	2814802.36	1	2814802.36	6095.44	.00	.97
Years Teaching	5475.11	4	1368.78	2.96	.02	.07
Education Level	1455.60	1	1455.60	3.15	.08	.02
Years Teach * Ed. Level	525.83	4	131.46	.28	.89	.01
Error	76656.85	166	461.79			
Total	4000516.00	176				
Corrected Total	85984.73	175				

The outlined rows in Table 4.19 show whether there were significant mean differences between the groups for the two independent variables, number of years teaching and highest education level completed, and for their interaction, Years Teach*Education Level. The main effect of Factor A—number of years teaching—addressed whether the number of years teaching had an effect on efficacy for literacy instruction. The main effect for Factor B—highest education level completed—addressed whether education level was related to teacher efficacy for literacy instruction. The interaction effect addressed whether the effects of number of years teaching depended on the education level of teachers.

The interaction effect between the number of years teaching and highest education level completed was not significant ($F(4,166)=.28$, $p=.89$), meaning that effects of number of years teaching are approximately the same regardless of the teachers' education level. Figure 4.9 depicts this result graphically. Visual inspection of Figure 4.9 shows that the effects of the number of years teaching are roughly the same regardless of education level.

Figure 4.9
Graphical Representation of the Test for Interaction



The test of the main effect of number of years teaching was significant ($F(4,166)=2.96$, $p<.05$), which was similar to the result obtained in the one-way ANOVA detailed in section 4.4.2.2.3. The test of the main effect of highest education level completed on TSELI scores was not significant ($F(1,166)=3.15$, $p=.08$) when tested simultaneously with the effect of number of years teaching. In section 4.4.2.3.3, the effect of highest education level completed was found to significantly impact teacher efficacy for literacy instruction when evaluated independently. That highest education level completed was not significant in the two-way ANOVA analysis is likely the result of shared variance between this factor and the number of years teaching.

Influence of Teacher Perceptions and Experiences on Efficacy for Literacy Instruction

Additional data were collected from study participants via a follow-up survey after teachers submitted their completed responses to the TSELI questionnaire. The follow-up survey contained four questions measuring teachers' perceptions of their undergraduate teacher-preparation program (measured on a 10-point low-to-high Likert scale), perceptions of their district-provided professional-development training (measured on a 10-point low-to-high Likert scale), the number of literacy-specific professional-development sessions the teacher attended during the applicable study year (1-2 sessions; 3-5 sessions; 6-8 sessions; 8+ sessions), and whether teachers ever had participated in a book club (yes or no). The surveys were provided to all 176 teachers who had completed the TSELI previously, but only 36 teachers (16 from District A, 9 from District B, and 11 from District C) completed the follow-up questionnaire.

Results from follow-up surveys.

Perception of undergraduate teacher preparation program.

Teachers were asked to rate the quality of their undergraduate teacher-preparation program on a 10-point scale ranging from low (negative perception) to high (positive perception). Responses ranged from 4 to 10. These data were recoded so that teachers responding with a 4 or 5 were classified into the *Low* group, those responding between 6 to 8 were classified into the *Mid* group, and those rating their teacher training program as a 9 or 10 were classified into the *High* group. Summary statistics for each of these groups are provided in Table 4.20. Table 4.21provides a summary of the other variables collected on the basis of the undergraduate-training program groups.

Table 4.20.
Summary Statistics Based on Perception of Quality of Teacher-Training Program

Group	N	Mean	SD	Min	Max	Avg. Response
Low	11	161.5	18.7	134	186	7.3
Mid	14	145.6	17.4	116	176	6.6
High	11	151.6	25.4	110	190	6.9
Missing	140	148.3	22.5	76	198	6.7

Table 4.21.
Crosstabulation of Variables Based on Undergraduate Training Perception Groups

Group	N	Avg. TSELI score	District PD rating	Number of District PD literacy session attended				Book Club	
				1-2	3-5	6-8	8+	Yes	No
Low	11	161.5	7.4	2	2	5	2	4	7
Mid	14	145.6	7.3	7	4	2	1	2	12
High	11	151.6	8.0	0	4	4	3	5	6

Table 4.20 reveals that teachers with the lowest perception of the quality of their undergraduate teacher-training program had the highest efficacy for literacy instruction scores on average. To determine whether this difference was statistically significant, a one-way ANOVA was conducted. Results are presented in Table 4.22, but differences in perceptions of undergraduate teacher training did not yield statistically significant TSELI scores ($F(2,35)=1.85$, $p=.17$).

Table 4.22.
One-way ANOVA: TSELI Score by Undergraduate Training Perceptions

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1560.52	2	780.26	1.85	.17
Within Groups	13904.70	33	421.35		
Total	15465.22	35			

Perceptions of the quality of district professional-development sessions.

Teachers were asked to rate the quality of their district's professional-development sessions on a 10-point scale ranging from low (negative perception) to high (positive perception). Responses ranged from 2 to 10. These data were recoded so that teachers responding with ratings of 2 to 5 were classified into the *Low* group, those responding between 6 to 8 were classified into the *Mid* group, and those rating their teacher-training program as a 9 or 10 were classified into the *High* group. Summary statistics for each of these groups are provided in Table 4.23. Table 4.24 provides a summary of the other variables collected on the basis of the district professional-development session groups.

Table 4.23.
Summary Statistics Based on Perception of Quality of District PD Sessions

Group	N	Mean	SD	Min	Max	Avg. Response
Low	5	150.0	19.6	134	176	6.8
Mid	21	154.1	19.5	116	190	7.0
High	10	149.6	26.2	110	186	6.8
Missing	140	148.3	22.5	76	198	6.7

Table 4.24.
Crosstabulation of Variables Based on District PD Perception Groups

Group	N	Avg. TSELI score	Undergrad rating	Number of District PD literacy session attended				Book Club	
				1-2	3-5	6-8	8+	Yes	No
Low	5	150.0	6.4	0	2	3	0	1	4
Mid	21	154.1	7.2	2	8	5	6	6	15
High	10	149.6	7.4	1	3	5	1	4	6

From Table 4.23, teachers with the highest ratings of the quality of their district professional-development sessions for literacy had the lowest average TSELI scores.

Frequency of professional-development sessions attended.

In the follow-up survey, teachers provided information regarding the number of literacy professional-development sessions they attended during the applicable study year. Two teachers attended 1-2 sessions, 14 attended between 3-5 sessions, 13 attended between 6-8 sessions, and seven participated in more than eight professional-development sessions during the year. Table 4.25 provides TSELI summary statistics for each of these groups.

Table 4.25.
Summary Statistics Based on Participation in Professional-Development Sessions

Frequency	N	Mean	SD	Min	Max	Avg. Response
1-2 sessions	2	183.0	4.2	180	186	8.3
3-5 sessions	14	142.9	24.9	110	182	6.5
6-8 sessions	13	156.6	17.7	134	196	7.1
8+ sessions	7	150.6	10.8	134	166	6.9
Missing	140	148.3	22.5	76	198	6.7

Table 4.26.
Crosstabulation of Variables Based on District PD Sessions Attended

Frequency	N	Avg. TSELI score	Undergrad rating	District PD rating	Book Club	
					Yes	No
1-2 sessions	2	183.0	4.5	8.0	0	2
3-5 sessions	14	142.9	7.5	7.4	5	9
6-8 sessions	13	156.6	6.9	7.6	5	8
8+ sessions	7	150.6	7.7	8.0	1	6

Participation in a book club.

Teachers were asked in the follow-up survey about personal enrichment experiences related to literacy, specifically whether they had participated in a book club. Eleven teachers (31%) indicated that they did participate in a book club while 25 (69%) responded that they did not. Teachers participating in a book club had a slightly higher TSELI average (153.1, *SD*=25.6) than those not participating in a book club (151.9, *SD*=19.3). An independent samples t-test was conducted to evaluate whether this difference was significant, and results indicated that it was not ($t(34)=.152$, $p=.88$).

Research question 3.

What is the influence of classroom-specific characteristics such as gender, English language learner status, special-education status, lunch status, and the ethnic makeup of students on teacher sense of efficacy for literacy instruction?

This research question addressed the effect of student characteristics on teacher efficacy for literacy instruction, specifically focusing on the context of individual teachers' classrooms through an analysis of student demographic variables aggregated at the class level. The effect of classroom contextual variables on efficacy for literacy instruction was assessed using ANOVA procedures to determine whether statistically significant differences existed between the means of TSELI scores for various groups based on student demographic characteristics such as gender, English Language Learner status, special-education status, lunch status, and ethnicity of students. For this series of analyses, three levels of each independent variable were created: Low, Mid, and High. Utilizing cumulative frequency distributions, the cut points for the three groups were arbitrarily set at the 33rd and 66th percentiles. In this manner, the *Low* group for each independent variable of interest included classrooms that fell between the 1st and 33rd percentile; the *Mid* group included classrooms that fell between the 34th and 66th percentile; and the *High* group consisted of classrooms that fell between the 67th and 99th percentile for the particular variable of interest. For the current series of analyses, only teacher and student data that could be matched were considered (n=138). For the respective analyses, the ANOVA procedure tested the null hypothesis:

$$H_0\colon \mu_L = \mu_M = \mu_H$$

where μ_k was the mean TSELI score for teachers in the Low, Mid, and High frequency groups for the particular classroom contextual variable of interest.

Influence of student gender on TSELI scores.

A one-way ANOVA was conducted to assess whether differences existed between TSELI scores for teachers whose classrooms varied in terms of the gender of students. For this question, the focal group was the total percentage of males in classrooms. Given that the gender variable only has two levels (male or female), the analysis only needed to be carried out on one of the two gender groups, as the results of the statistical tests would be identical for both groups given the inverse nature of the relationship. Of the 3,814 students in the study sample, 1,866 (49%) were female, and 1,948 (51%) were male.

Testing of assumptions.

The assumptions of ANOVA (detailed in section 4.4.1.1) were tested prior to conducting this analysis. No violations of normality or homogeneity of variances were observed.

Results of analysis.

Table 4.27 provides the summary statistics from the ANOVA analysis, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for the Low, Mid, and High classrooms and all groups combined. Figure 4.10 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.27
Summary statistics: TSELI Scores by Gender of Class

% of males	N*	Mean TSELI	SD	Std. Error	95% CI		Min	Max	Avg. Response
					Lower	Upper			
Low	46	150.8	21.3	3.1	144.5	157.1	108	192	6.9
Mid	46	151.0	22.1	3.3	144.5	157.6	110	198	6.9
High	46	151.5	19.8	2.9	145.6	157.4	112	192	6.9
Total	138	151.1	21.0	1.8	147.6	154.6	108	198	6.9

*Number of classrooms

Figure 4.10.
Average TSELI Score: Gender Composition of Classroom

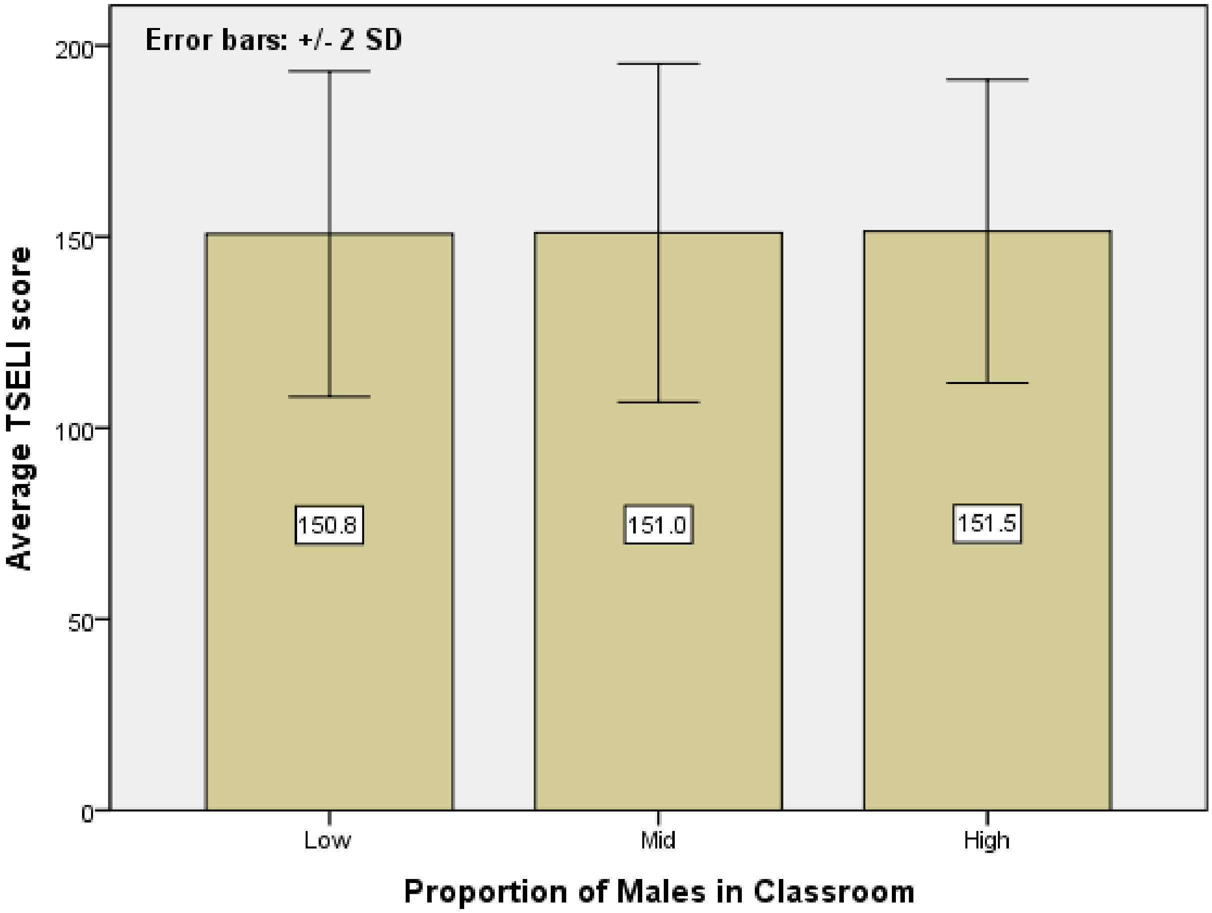


Table 4.28 displays the results from the one-way ANOVA. There was not a statistically significant difference between TSELI scores for teachers whose classrooms varied in terms of the proportion of male and female students as determined by the one-way ANOVA (F(2,137) = .01, p = .99). The effect size (□^{2}) was 0, indicating that student gender did not account for any variation in TSELI scores.

Table 4.28.
One-way ANOVA: TSELI Scores by Gender of Class

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	11.36	2	5.68	.01	.99
Within Groups	60133.22	135	445.43		
Total	60144.58	137			

The one-way ANOVA was conducted to compare the effect of student gender on TSELI scores. Classrooms were divided into three groups based on the proportion of males in each class: Low, Mid, and High. Results from the one-way ANOVA indicated that teachers' sense of efficacy for literacy instruction did not significantly differ based on student gender.

Influence of English language learners on TSELI scores.

For this analysis, a one-way ANOVA was used to determine whether classrooms with varying levels of students classified as English Language Learners (ELLs) influenced TSELI scores. In the study sample, 329 (9%) of the 3,814 students were classified as ELLs.

Testing of assumptions.

The assumptions of ANOVA (detailed in section 4.4.1.1) were evaluated prior to conducting this analysis. No violations of normality or homogeneity of variances were observed.

Results of analysis.

Table 4.29 provides the summary statistics from the ANOVA analysis, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for the Low, Mid, and High classrooms and all groups combined. Figure 4.11 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.29
Summary statistics: TSELI Scores by Classroom ELL Status

% of ELL	**N***	**Mean TSELI**	**SD**	**Std. Error**	**95% CI**		**Min**	**Max**	**Avg. Response**
					Lower	**Upper**			
Low	52	151.7	22.8	3.2	145.3	158.0	108	192	6.9
Mid	45	151.4	21.1	3.1	145.0	157.7	110	198	6.9
High	41	150.1	18.7	2.9	144.2	156.0	112	186	6.8
Total	138	151.1	21.0	1.8	147.6	154.6	108	198	6.9

*Number of classrooms

Figure 4.11.
Average TSELI Score: ELL Composition of Classroom

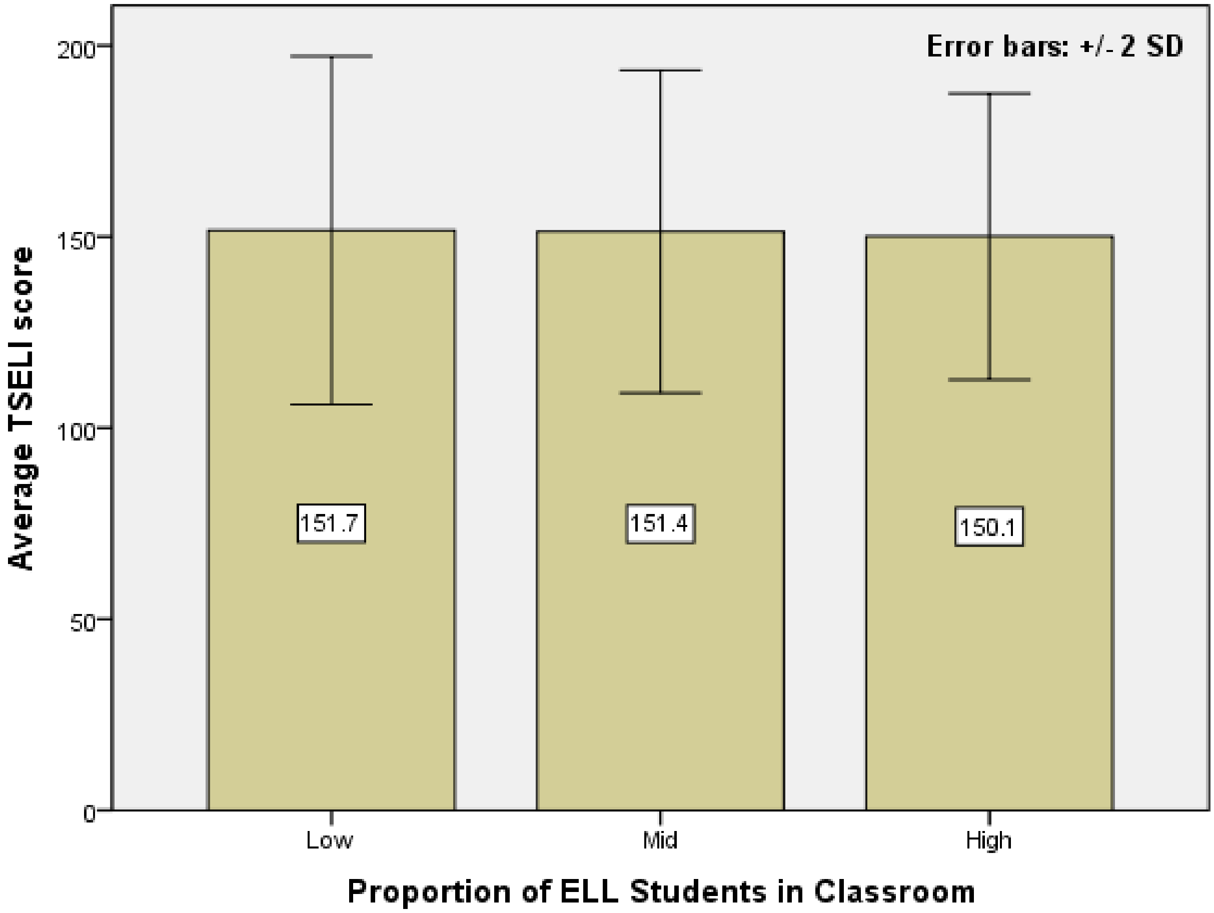


Table 4.30 displays the results from the one-way ANOVA. There was not a statistically significant difference between TSELI scores for teachers whose classrooms varied in terms of the proportion of students classified as English Language Learners ($F(2,137) = .07$, $p = .93$). The effect size (□^{2}) was 0, indicating that student ELL status did not account for any variation in TSELI scores.

Table 4.30.
One-way ANOVA: TSELI Scores by Classroom ELL Status

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	60.62	2	30.31	.07	.93
Within Groups	60083.96	135	445.07		
Total	60144.58	137			

The one-way ANOVA was conducted to compare the effect of student ELL status on TSELI scores. Classrooms were divided into three groups based on the proportion of students classified as English Language Learners in each class: Low, Mid, and High. Results from the one-way ANOVA indicated that teachers' sense of efficacy for literacy instruction did not significantly differ based on student ELL status.

Influence of students receiving special-education services on TSELI scores.

A one-way ANOVA was employed to assess the influence of Special-Education (SPED) students on teachers' efficacy for literacy instruction. This analysis evaluated classrooms with varying levels of SPED students in order to determine the impact of this classroom contextual variable on TSELI scores. In the study sample, 448 (12%) of the 3,814 students were classified as receiving special-education services.

Testing of assumptions.

The assumptions of ANOVA (detailed in section 4.4.1.1) were evaluated prior to conducting this analysis. No violations of normality or homogeneity of variances were observed.

Results of analysis.

Table 4.31 provides the summary statistics from the ANOVA analysis, including the mean, standard deviation ,and 95% confidence intervals for the dependent variable (TSELI score) for classrooms with Low, Mid, and High proportions of SPED students and the total for all groups combined. Figure 4.12 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.31.
Summary statistics: TSELI Scores by Classroom SPED Status

% of SPED	N*	Mean TSELI	SD	Std. Error	95% CI Lower	95% CI Upper	Min	Max	Avg. Response
Low	47	149.8	17.6	2.6	144.7	155.0	112	184	6.8
Mid	49	150.7	23.7	3.4	143.9	157.5	108	198	6.9
High	42	153.0	21.3	3.3	146.3	159.7	112	192	7.0
Total	138	151.1	21.0	1.8	147.6	154.6	108	198	6.9

*Number of classrooms

Figure 4.12.
Average TSELI score: SPED Composition of Classroom

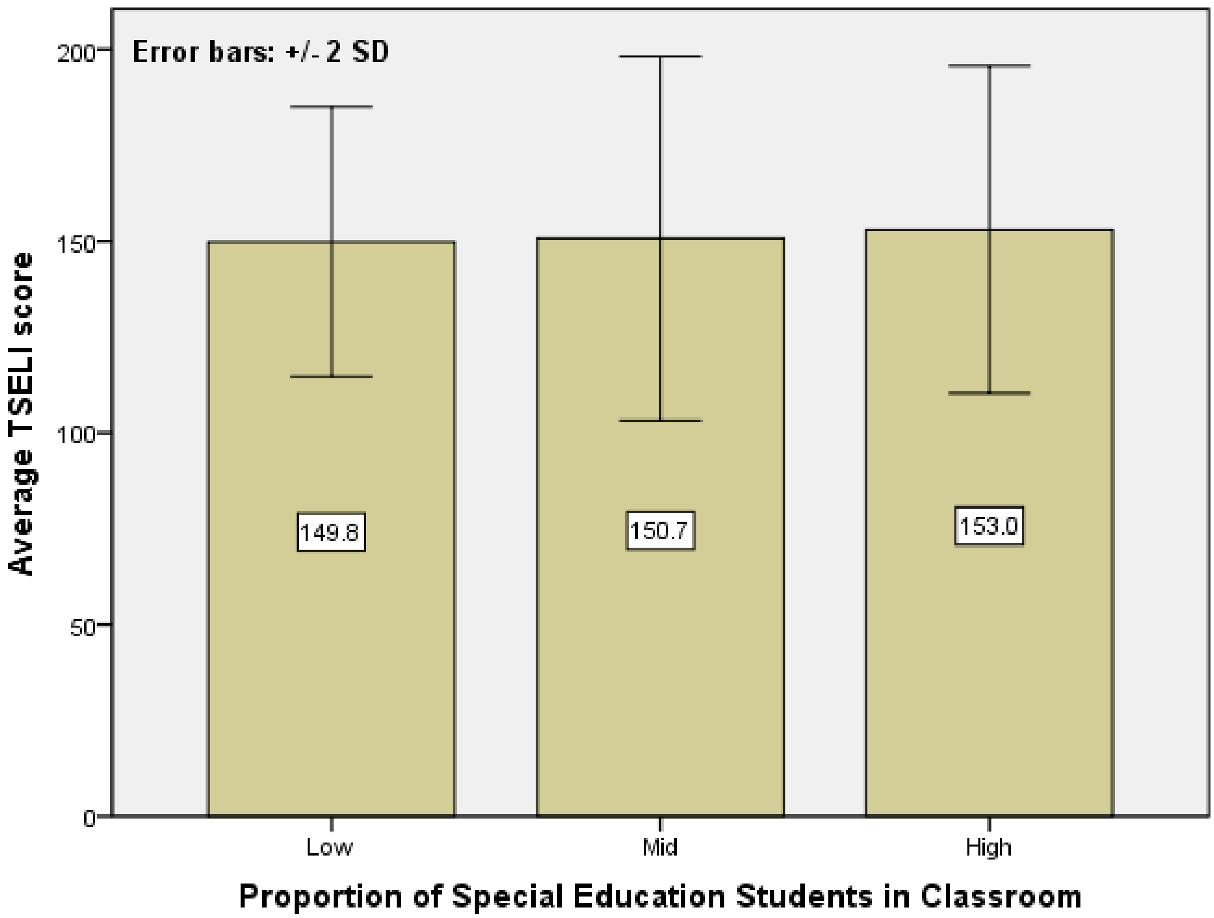


Table 4.32 displays the results from the one-way ANOVA. While teachers in classrooms with higher proportions of SPED students reported higher levels of efficacy for literacy instruction, this difference was not statistically significant ($F(2,137) = .27$, $p = .77$). The effect

size (□^{2}) was 0, indicating that the classroom composition in terms of Special-Education students did not account for any variation in TSELI scores.

Table 4.32.
One-way ANOVA: TSELI Scores by Classroom SPED Status

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	235.53	2	117.77	.27	.77
Within Groups	59909.05	135	443.77		
Total	60144.58	137			

The one-way ANOVA was conducted to compare the effect of student SPED status on TSELI scores. Classrooms were divided into three groups based on the proportion of students receiving Special-Education services in each class: Low, Mid, and High. Results from the one-way ANOVA indicated that teachers' sense of efficacy for literacy instruction did not significantly differ based on student SPED status.

Influence of student ethnicity on TSELI scores.

This analysis addressed the influence of student ethnicity on TSELI scores via a one-way ANOVA. Detailed information on ethnicity was provided in the student data file supplied by each of the participating districts in the study. For this analysis, ethnicity was recoded to reflect two levels of the variable: Caucasian and Non-Caucasian students. The study sample consisted of 2,367 (62%) Caucasian students and 1,447 (38%) non-Caucasian students. Percentages of Caucasian students associated with each teacher (n=138) in the analysis sample were ascertained and categorized into three levels with cut points at the 33rd and 66th percentile: Low, Mid, and High. This analysis evaluated classrooms with varying levels of Caucasian students in order to determine the impact of ethnicity on TSELI scores.

Testing of assumptions.

The assumptions of ANOVA (detailed in section 4.4.1.1) were evaluated prior to conducting this analysis. No violations of normality or homogeneity of variances were observed.

Results of analysis.

Table 4.33 provides the summary statistics from the ANOVA analysis, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for classrooms with Low, Mid, and High proportions of Caucasian students and the total for all groups combined. Figure 4.13 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.33.
Summary Statistics: TSELI Scores by Classroom Ethnicity

% of Caucasian	N*	Mean TSELI	SD	Std. Error	95% CI Lower	95% CI Upper	Min	Max	Avg. Response
Low	44	147.7	19.6	3.0	141.7	153.7	110	192	6.7
Mid	46	146.7	17.8	2.6	141.4	151.9	110	180	6.7
High	48	158.5	23.2	3.3	151.8	165.2	108	198	7.2
Total	138	151.1	21.0	1.8	147.6	154.6	108	198	6.9

*Number of classrooms

Figure 4.13. Average TSELI score: Ethnic Composition of Classroom

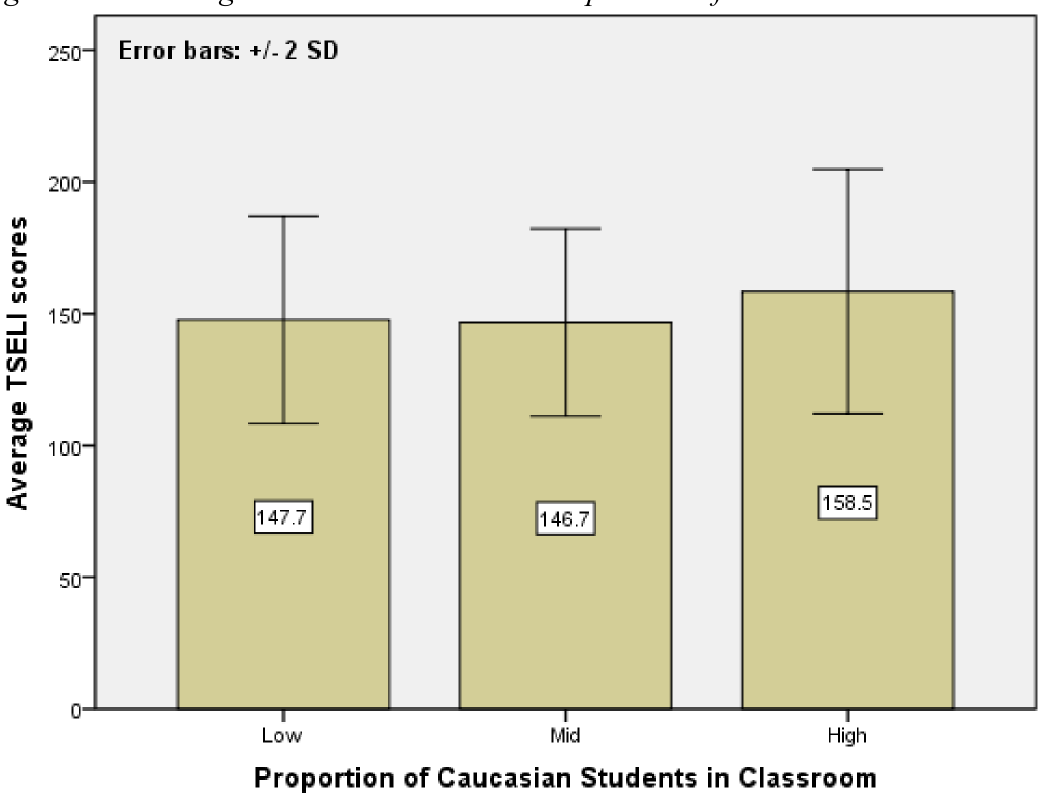


Table 4.34 displays the results from the one-way ANOVA. There was a statistically significant difference in TSELI scores between groups as determined by the one-way ANOVA ($F(2,137) = 4.88$, $p < .05$). The effect size ($\square^2$) was .07, indicating that student ethnicity accounted for 7% of the variance in TSELI scores.

Table 4.34.
One-way ANOVA: TSELI Scores by Classroom Ethnicity Status

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4052.60	2	2026.30	4.88	.01
Within Groups	56091.98	135	415.50		
Total	60144.58	137			

The one-way ANOVA was conducted to compare the effect of student ethnicity on TSELI scores. Results from the analysis indicated that teachers' sense of efficacy for literacy instruction significantly differed based on the number of Caucasian students in their classes. Given the significant difference resulting from the omnibus ANOVA, the Tukey HSD post-hoc test was used to confirm where differences occurred between groups. Table 4.35 shows the Tukey post-hoc test results for the different levels of Caucasian students.

Table 4.35.
Multiple Comparisons for Effect of Student Ethnicity

(I) % Caucasian students	(J) % Caucasian students	Mean Difference (I-J)	Std. Error	Sig.	95% CI for the Difference	
					Lower bound	Upper bound
Low	Mid	1.03	4.30	.97	-9.16	11.22
	High	-10.82*	4.25	.03	-20.90	-.74
Mid	Low	-1.03	4.30	.97	-11.22	9.16
	High	-11.85*	4.21	.02	-21.81	-1.88
High	Low	10.82*	4.25	.03	.74	20.90
	Mid	11.85*	4.21	.02	1.88	21.81

*The mean difference is significant at the 0.05 level

From the results of the Tukey post-hoc tests presented in Table 4.35, significant mean TSELI differences existed between classes with the highest proportion of Caucasian students versus the other two groups. In effect, teachers with more Caucasian students had higher levels of efficacy for literacy instruction.

Influence of lunch program status on TSELI scores.

This analysis addressed the influence of student participation the Free/Reduced lunch program on TSELI scores via a one-way ANOVA. For this analysis, lunch status was recoded to reflect two levels of the variable: students participating in the program (free or reduced lunch) and those not participating. In the study sample, 1,618 (42%) students participated in the free/reduced lunch program. This analysis evaluated classrooms with varying levels of students participating in the lunch program in order to determine the impact on TSELI scores.

Testing of assumptions.

The assumptions of ANOVA (detailed in section 4.4.1.1) were evaluated prior to conducting this analysis. No violations of normality or homogeneity of variances were observed.

Results of analysis.

Table 4.36 provides the summary statistics from the ANOVA analysis, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (TSELI score) for classrooms with Low, Mid, and High proportions of students participating in the lunch program and the total for all groups combined. Figure 4.14 provides a graphical representation of each group's TSELI mean along with standard-deviation bars (+/- 2 SD).

Table 4.36
Summary Statistics: TSELI Scores by Classroom Lunch Status

% of Lunch	N*	Mean TSELI	SD	Std. Error	95% CI		Min	Max	Avg. Response
					Lower	Upper			
Low	44	158.8	22.5	3.4	151.9	165.6	108	198	7.2
Mid	46	148.3	20.9	3.1	142.1	154.5	110	192	6.7
High	48	146.7	17.8	2.6	141.5	151.9	110	178	6.7
Total	138	151.1	21.0	1.8	147.6	154.6	108	198	6.9

*Number of classrooms

Figure 4.14.
Average TSELI Score: Lunch Program Composition of Classroom

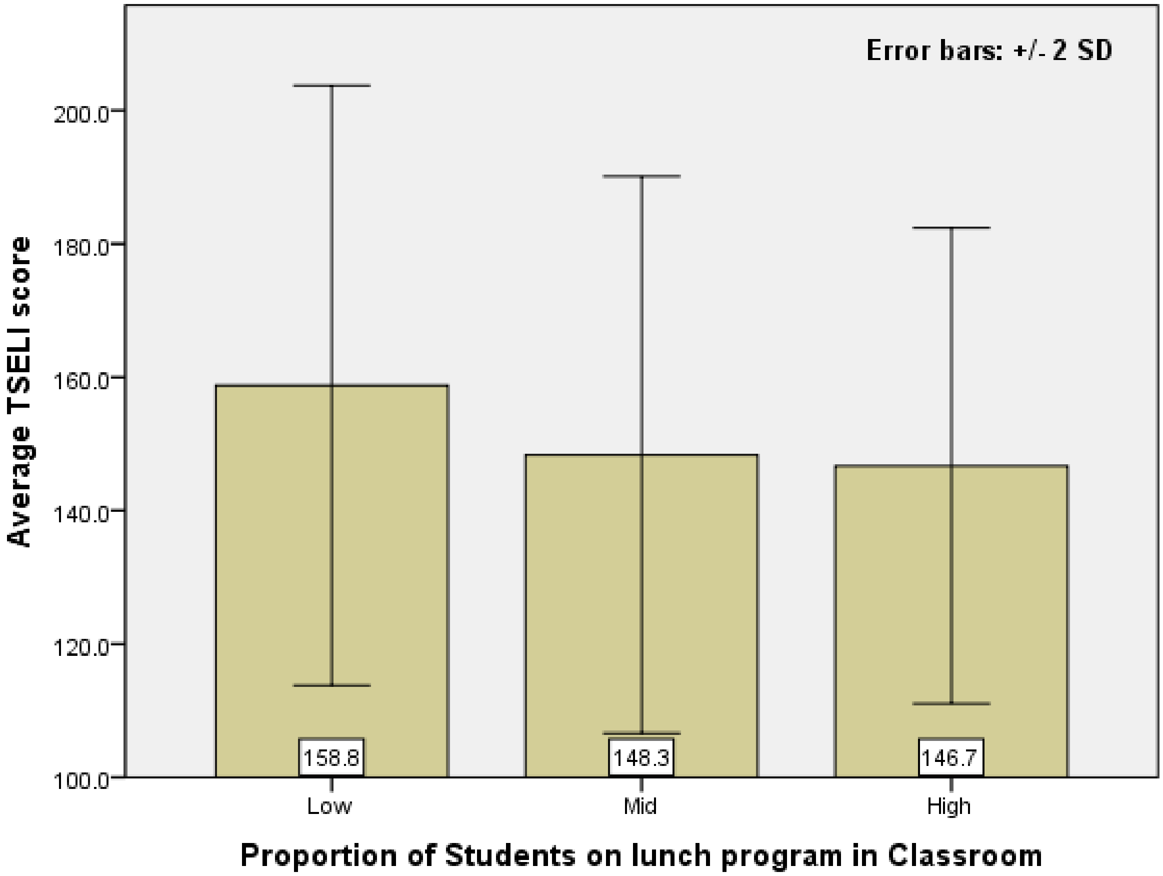


Table 4.37 displays the results from the one-way ANOVA. There was a statistically significant difference in TSELI scores between groups as determined by the one-way ANOVA ($F(2,137) = 4.63$, $p < .05$). The effect size (□^{2}) was .06, indicating that lunch-program status accounted for 6% of the variance in TSELI scores.

Table 4.37
One-way ANOVA: TSELI Scores by Classroom Lunch Status

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3864.50	2	1932.25	4.63	.01
Within Groups	56280.08	135	416.89		
Total	60144.58	137			

The one-way ANOVA was conducted to compare the effect of student participation in the Free/Reduced lunch program on TSELI scores. Results from the analysis indicated that teachers' sense of efficacy for literacy instruction significantly differed based on the number of students on free/reduced lunch in their classes. Given the significant difference resulting from the omnibus ANOVA, the Tukey HSD post-hoc test was used to confirm where differences occurred between groups. Table 4.38 shows the Tukey post-hoc test results for the different levels of students in the lunch program.

Table 4.38.
Multiple Comparisons for Effect of Lunch Status

(I) % Caucasian students	**(J) % Caucasian students**	**Mean Difference (I-J)**	**Std. Error**	**Sig.**	**95% CI for the Difference**	
					Lower bound	Upper bound
Low	Mid	10.42*	4.31	.04	.22	20.63
	High	12.06*	4.26	.01	1.97	22.16
Mid	Low	-10.42*	4.31	.04	-20.63	-.22
	High	1.64	4.21	.92	-8.34	11.62
High	Low	-12.06*	4.26	.01	-22.16	-1.97
	Mid	-1.64	4.21	.92	-11.62	8.34

*The mean difference is significant at the 0.05 level

From the results of the Tukey post-hoc tests presented in Table 4.38, significant mean TSELI differences existed between classes with the lowest proportion of students participating in the free/reduced lunch program versus the other two groups. In effect, teachers with fewer students participating in the program had higher levels of efficacy for literacy instruction.

Research question 4.

What is the impact of teachers' sense of efficacy for literacy instruction on student reading achievement?

Whereas the previous research questions addressed in this study examined the impact of certain teacher and student variables on efficacy for literacy instruction, this question addressed the relationship and influence of the construct on student reading achievement through two analytic approaches, analysis of variance (ANOVA) and hierarchical linear modeling (HLM).

Student reading achievement was measured by scores on the Kansas Reading Assessment, a criterion-referenced assessment administered by the state of Kansas as part of its statewide testing program mandated by the federal Elementary and Secondary Education Act (ESEA). Each grade-level assessment was developed to measure the Kansas curricular standards in reading. Technical documentation of the Kansas Reading Assessment was delineated by Poggio, Yang, Irwin, Glasnapp, and Poggio (2007).

To address this impact of efficacy for literacy instruction on student reading achievement, two analytic approaches were employed. The first approach utilized analysis of variance techniques to document the relationship between reading achievement at various levels of teacher efficacy for literacy instruction. The second approach applied hierarchical modeling to account for the nested nature of the study data in order to evaluate the impact of teacher efficacy for literacy instruction on student reading achievement.

Relationship of efficacy for literacy instruction on reading achievement: ANOVA

A one-way ANOVA was conducted to determine whether differences existed in student reading achievement at various levels of teacher efficacy for literacy instruction. Teacher TSELI scores were recoded to reflect three groups relative to the distribution of TSELI scores in the study sample: *Low* teacher efficacy, *Mid* teacher efficacy, and *High* teacher efficacy. The levels were calculated based on the mean ($\bar{X}$= 151.1) and standard deviation (SD=21.0) for the overall study sample (n=138). The *Low* efficacy group included teachers whose TSELI scores were greater than one standard deviation below the overall group mean (TSELI≤130). The *Mid* efficacy group included all teachers whose TSELI scores were within one standard deviation above and below the group mean (131<TSELI<172). The High efficacy group included teachers whose TSEL scores were greater than one standard deviation above the group mean (TSELI≥173).

It should be noted that this analysis did not separate student reading achievement on the basis of grade level. While each grade level reading assessment measures unique content, as a criterion-referenced assessment, the total score has the same meaning and interpretation regardless of grade level.

Testing of assumptions.

The assumptions of ANOVA were tested prior to executing this analysis. No violations of the normality were observed, but the assumption of equal variances was not met. Given the violation of this assumption, the Welch Robust Test of Equality of Means was used to determine whether a significant difference existed between reading achievement scores at varying levels of teacher efficacy for literacy instruction.

Results of analysis.

Table 4.39 provides the summary statistics from the ANOVA analysis, including the mean, standard deviation, and 95% confidence intervals for the dependent variable (student reading achievement score) for the TSELI score groups as well as for the groups combined. Figure 4.15 provides a graphical representation of each group's reading achievement mean along with standard-deviation bars (+/- 2 SD).

Table 4.39.
Summary Statistics: Reading Scores by Teacher TSELI Level

TSELI level	N*	Mean	SD	Std. Error	95% CI		Min	Max
					Lower	Upper		
Low	682	77.8	14.2	.5	76.8	78.9	26	100
Mid	2464	79.8	14.1	.3	79.2	80.3	24	100
High	668	81.8	12.9	.5	80.8	82.8	23	100
Total	3814	79.8	13.9	.2	79.4	80.2	23	100

*Number of students in each group

Figure 4.15.
Average Reading Score for Teacher Efficacy Levels

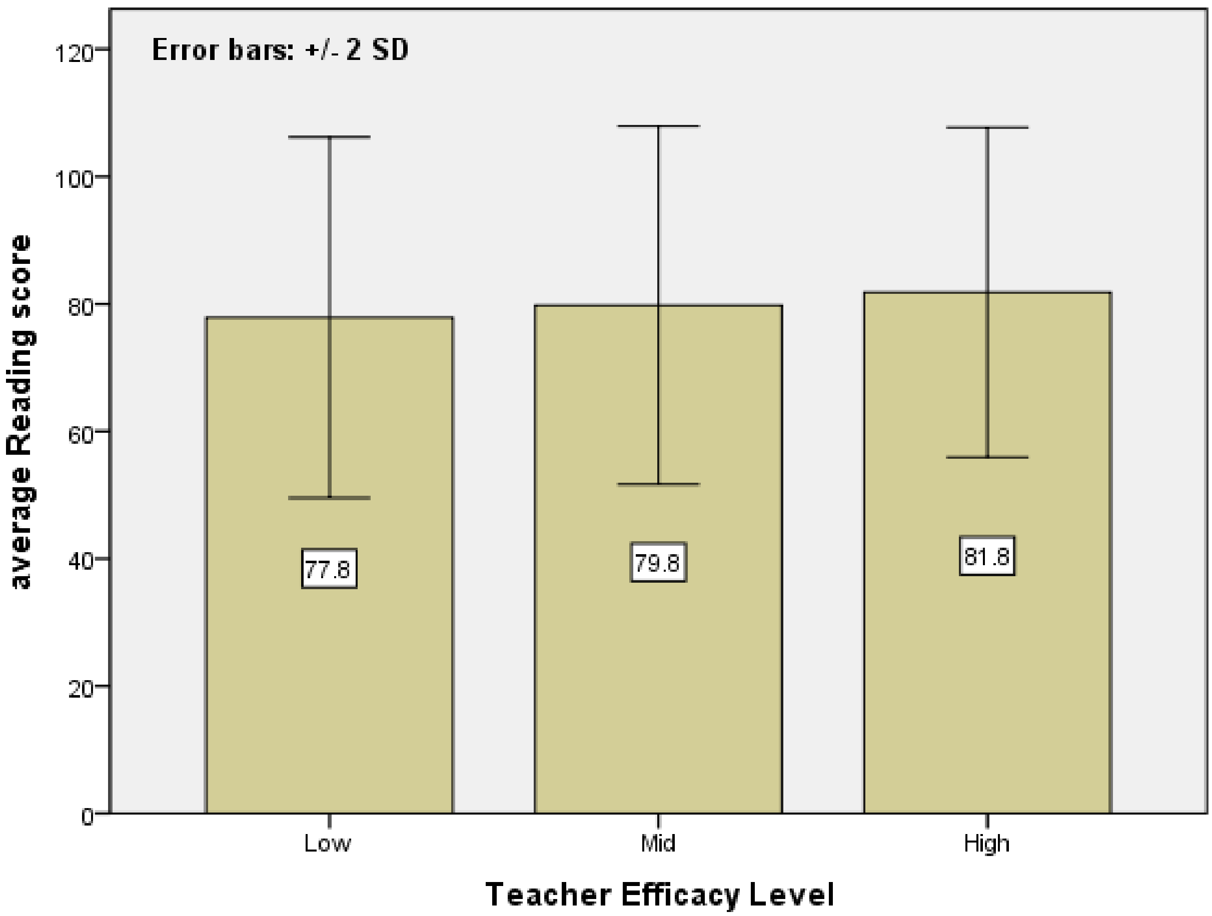


Table 4.40 displays the results from the ANOVA analysis. As the assumption of equal variances was violated, the Welch Robust Test of Equality of Means also was consulted to verify the ANOVA test result. There was a statistically significant difference in reading achievement between groups as determined by the one-way ANOVA ($F(2,3813) = 13.75$, $p < .01$). The effect size, however, was small ($\eta^2 = .01$), indicating that levels of teacher efficacy for literacy instruction accounted for 1% of the variation in reading achievement.

Table 4.40.
One-way ANOVA: Reading Scores by Teacher TSELI Level

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5298.53	2	2649.27	13.75	.00
Within Groups	734471.72	3811	192.72		
Total	739770.25	3813			
		Statistic	**df1**	**df2**	**Sig.**
Welch Test		14.51	2	1318.90	.00

The one-way ANOVA was conducted to compare the effect of levels of teacher efficacy for literacy instruction on student reading achievement. Results from the analysis indicated that average reading achievement significantly differed based on the level of teacher efficacy. Given the significant difference resulting from the omnibus ANOVA, the Tukey HSD post-hoc test was used to confirm where differences occurred between groups. Table 4.41 shows the Tukey post-hoc test results for the different levels of teacher efficacy for literacy instruction.

Table 4.41.
Multiple Comparisons for Effect of Teacher Efficacy

(I) TSELI score	(J) TSELI score	Mean Difference (I-J)	Std. Error	Sig.	95% CI for the Difference	
					Lower bound	Upper bound
Low	Mid	-1.94*	.60	.00	-3.35	-.53
	High	-3.96*	.76	.00	-5.73	-2.19
Mid	Low	1.94*	.60	.00	.53	3.35
	High	-2.02*	.61	.00	-3.44	-.60
High	Low	3.96*	.76	.00	2.19	5.73
	Mid	2.02*	.61	.00	.60	3.44

*The mean difference is significant at the 0.01 level

From the results of the Tukey post-hoc tests presented in Table 4.41, significant mean TSELI differences in student reading achievement existed among all levels of teacher efficacy for literacy instruction.

Impact of efficacy for literacy instruction on reading achievement: HLM.

To examine the impact of efficacy for literacy instruction on student reading achievement, hierarchical linear modeling (HLM) was conducted. Hierarchical linear models are designed to analyze data in which lower-level units of analysis are nested in higher-level units of analysis. For example, students are nested within classrooms, which are nested within schools. When data are nested, student variables are likely to be less independent. For example, it is often the case that the reading achievement of students within a particular class is more similar than in a random sample of students drawn across different classrooms. This is because students in the same classroom share a common curriculum, common textbooks, and a common teacher, for example. HLM is preferable to Ordinary Least Squares (OLS) regression, as traditional regression models treat all students as independent, potentially obscuring differences between higher-level units (such as classes). A major concern when using OLS regression to estimate relationships on nested data is that the estimated standard errors will be too small (negatively biased), leading to an overestimation of the statistical significance of regression coefficients. In contrast to regression models, hierarchical models first make predictions about higher-level parameters (for example, class level), which then are used to make predictions about lower-level parameters (student level).

For this analysis, a series of hierarchical linear models with two levels were constructed to assess the impact of teacher efficacy for literacy instruction on student reading achievement. A two-level model consists of level-2 units in which the level-1 units are nested. For example, students, the level-1 unit of analysis, were nested within classrooms, the level-2 unit. In this analysis, the relationship between TSELI scores and reading achievement for students of 138 classroom teachers was analyzed. By utilizing a hierarchical model, the goal of this analysis was to determine the impact of teacher efficacy on student reading achievement, and if a relationship did exist, to determine whether students from different classrooms showed systematic differences in the strength of the relationship between teacher efficacy and reading achievement.

Assumptions of HLM.

Assumptions for employing hierarchical models are similar to those of OLS regression, wexcept that they must be examined at each level. For this analysis, the following was assumed (Bryk and Raudenbush, 1992):

- The error term of each level-1 unit should have a mean of zero, and the residuals should be normally distributed. For example, if the level-1 units are students and level-2 units are classrooms, then the mean of the error within each classroom should be zero, the residuals should be normally distributed, and all classrooms should have variances equal to the other classrooms in the sample.

- Level-1 predictors are independent of the level-1 error term. That is, the covariance between the level-1 predictors and the error term should equal zero.

- Level-2 error terms each have a mean of zero and adhere to a multivariate normal distribution.

- Level-2 predictors are independent of all level-2 error terms. Thus all variables in the second level of the model are not related to any of the error terms on that level of the model, including the error term for the level-1 intercept and the error term for the slopes of level-1 variables.

- The level-1 error terms are independent of level-2 error terms. That is, there is no relationship between the error term at level-1 and the error term in the level-2 equation for the level-1 intercept, or the error term in any of the equations used to estimate the slopes of level-1 variables.

All level-1 and level-2 residuals (error terms) were examined for potential violations of normality, linearity, and homogeneity of variance in accordance with the criteria listed above. For error terms at both levels, plots of residuals were analyzed as well as the skewness and kurtosis of the residuals. Examination of these numerical and graphical indices did not reveal any violations of the HLM assumptions. Both level-1 and level-2 residuals were distributed normally.

Additionally, normality, linearity, and heteroscedasticity were further examined by plotting the level-2 residuals against the predicted values for student reading achievement. Once again, no violations of the assumptions were noted.

Results of analysis.

The first step in conducting an HLM analysis was to determine the extent to which observations within classes were correlated, that is, the degree of non-independence among observations in a sample. An unconditional (also referred to as a "null") model was first run to partition the variance in the dependent variable, in this case reading achievement, into within- and between-classroom components. The unconditional model includes no predictor variables at level-1 or level-2 and indicates the average class-level reading achievement for the entire sample. After the unconditional model, level-2 main effects were added to the model followed by the inclusion of the student level (level-1) covariates. Table 3.12 includes a general overview of the structure of the hierarchical models constructed for this analysis.

The unconditional model. The unconditional model includes no predictors at either level-1 or level-2 and is intended to be a baseline model that serves two purposes. First, it partitions variation in the dependent variable into two components, between- and within-classes. Second, the unconditional model serves as a baseline against which all subsequent models can be evaluated and compared. This model is represented as

$$Y_{ij} = \beta_{0j} + r_{ij}$$

and was used to predict reading achievement within each class with just one class level (level-1) paramenter, the intercept β_{0j}, which is the mean reading achievement outcome for class j. The level-2 model is represented as

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

where γ_{00} represents the grand mean outcomes (overall reading achievement) and u_{0j} is the random effect associated with unit j. By substituting the level-2 model into the level-1 model

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$$

it is easier to see how the variance is being partitioned. That is, an individual student's reading score can be composed into the overall mean score (γ_{00}), how much the student's score differs from the overall mean score (u_{0j}), and how much the individual student's score differs from the class average score (r_{ij}). It is the partitioning of the residual or "random" component of the model into class and student level components that make this a multilevel model.

The unconditional model partitions variation in the dependent variable of reading achievement (Y_{ij}) into two components: between classes ($Var(u_{0j}) = \tau_{00}$) and within classes ($Var(r_{ij}) = \sigma^2$). The unconditional model is fit to data to estimate both the fixed effect, γ_{00}, or the average reading achievement for the overall sample, and the two random effects, τ_{00} indicating the variability in class means and σ^2 indicating the variability in reading achievement within classes. Results from the unconditional model are provided in Table 4.42.

Table 4.42
Fixed and Random Effects for the 2-Level Unconditional HLM Model

Fixed Effects	**Coefficients**	**SE**	***t*-ratio**	***p***
Intercept, γ_{00}	79.86	.63	127.00	<.001
Random Effects	**Variance components**	**SD**	**$\chi 2$**	***p***
Intercept, τ_{00}	48.60	6.97	1693.75	<.001
Residual, σ^2	137.74	11.74		
Model Fit	**Value**	**Parameters**		
Deviance	29919.73	2		

*Data represent 3,814 students and 138 teachers

For the unconditional model, the estimated value of $\tau_{00} = 48.60$ and the estimated value of $\sigma^2 = 137.74$. Hypothesis tests for the between-classes variation component indicates that it is significantly different than zero, suggesting that classes do differ in their average reading achievement. Further, the results indicate that there is modest variation among students within classes, since the variance component within classes (σ^2) is nearly three times the size of the variance component between classes (τ_{00}).

Another way to consider the proportion of the total variance that is between classes is to estimate the Intraclass Correlation Coefficient (ICC), which is equal to

$$\rho = \tau_{00}/(\tau_{00} + \sigma^2)$$

which in this model is estimated to be

$$48.60/(48.60 + 137.74) = .26$$

Thus for this sample, 26% of the variation in student's reading achievement is between classes. Given this ICC value and the significance test results, it is evident that there is significant variation in student reading achievement that could be explained by adding class- and student-level variables in the model. In other words, it is justified to continue to use HLM to explain the variation in student reading achievement.

Adding class-level (level-2) predictors. The second family of models constructed in this analysis contained the class-level main effects of efficacy for literacy instruction score, number of years teaching, education level, and district.

The primary approach to model building was to construct a series of models with each of the level-2 predictors entered individually. After assessing the relative impact of each class-level predictor, combinations of variables were entered to determine the "best" main-effects model based on level-2 predictors. Results from the entry of level-2 variables into the model individually are presented in Table 4.43.

Table 4.43.
Fixed and Random Effects for the Level-2 HLM Models

Fixed Effects	**Coefficients**	**SE**	***t*-ratio**	***p***
Intercept, γ_{00}	85.11	1.37	62.26	<.001
District, γ_{01}	-2.79	.66	-4.21	<.001
Random Effects	**Variance components**	**SD**	**$\chi 2$**	***p***
Intercept, τ_{00}	43.08	5.56	1641.57	<.001
Residual, σ^2	137.74	11.74		
Model Fit	**Value**	**Parameters**	**Variance Accounted for**	
Deviance	29903.67	2	11%	
Fixed Effects	**Coefficients**	**SE**	***t*-ratio**	***p***
Intercept, γ_{00}	77.25	1.43	54.19	<.001
Years Teach, γ_{01}	.89	.43	2.04	<.05
Random Effects	**Variance components**	**SD**	**$\chi 2$**	***p***
Intercept, τ_{00}	47.46	6.89	1634.76	<.001
Residual, σ^2	137.74	11.74		
Model Fit	**Value**	**Parameters**	**Variance Accounted for**	
Deviance	29917.44	2	2%	
Fixed Effects	**Coefficients**	**SE**	***t*-ratio**	***p***
Intercept, γ_{00}	74.72	1.99	37.52	<.001
Ed. Level, γ_{01}	1.71	.63	2.71	<.01
Random Effects	**Variance components**	**SD**	**$\chi 2$**	***p***
Intercept, τ_{00}	46.26	6.80	1621.50	<.001
Residual, σ^2	137.74	11.74		
Model Fit	**Value**	**Parameters**	**Variance Accounted for**	
Deviance	29913.43	2	5%	
Fixed Effects	**Coefficients**	**SE**	***t*-ratio**	***p***
Intercept, γ_{00}	71.24	4.54	15.71	<.001
TSELI, γ_{01}	.06	.03	2.18	<.05
Random Effects	**Variance components**	**SD**	**$\chi 2$**	***p***
Intercept, τ_{00}	47.54	6.90	1641.06	<.001
Residual, σ^2	137.74	11.74		
Model Fit	**Value**	**Parameters**	**Variance Accounted for**	
Deviance	29923.10	2	2%	

The regression coefficients for each of the level-2 predictors was significant, yet the variables generally did not explain much variation in student reading achievement with the exception of the District variable (11%).

The overall focus of this research question addressed the impact of teacher efficacy for literacy instruction on student reading achievement, so the results for TSELI entered into the model individually is of particular note. While the regression coefficient for TSELI was significant, efficacy for literacy instruction did not account for much variation in student reading achievement (2%). The regression coefficient for TSELI was .06, meaning that for each one-point increase in TSELI scores, student reading achievement is predicted to increase by .06 points.

When entered into the hierarchical model with the other level-2 variables, efficacy for literacy instruction remained significant as a predictor when paired with Education Level and District (p<.05), but was not significant when entered into the model with Years Teaching (p=12). When TSELI, District, and Education Level were entered together as class-level predictors of student reading achievement, the effect of Education Level was not significant. Therefore the final model of level-2 predictors included both TSELI and District. Summary statistics for the retained model are provided in Table 4.44. This model explained 13% of the variation in student reading achievement.

Table 4.44.
Fixed and Random Effects for the Retained Level-2 HLM Model

Fixed Effects	**Coefficients**	**SE**	***t*-ratio**	***p***
Intercept, γ_{00}	76.82	3.95	19.43	<.001
District, γ_{01}	-2.76	.62	-4.42	<.001
TSELI, γ_{02}	.05	.03	2.09	<.05
Random Effects	**Variance components**	**SD**	**$\chi 2$**	***p***
Intercept, τ_{00}	42.11	6.49	1590.67	<.001
Residual, σ^2	137.74	11.74		
Model Fit	**Value**	**Parameters**	**Variance Accounted for**	
Deviance	29903.42	2	13%	

Adding student-level (level-1) predictors. The third family of models considered the student-level (Level-1) covariates of ethnicity, English Language Learner status, Special-Education status, free/reduced lunch program status, grade, and gender. Variations of model building found grade and gender not to be significant, so these variables were not retained. Additionally, potential interactions between all level-1 predictors and the main effect of efficacy for literacy instruction (TSELI score) were evaluated; no significant interactions were found. Summary statistics for the full, final hierarchical linear model of student reading achievement are presented in Table 4.45.

Table 4.45.
Hierarchical Linear Model of Best Fit for Student Reading Achievement

Fixed Effects	**Coefficients**	**SE**	***t*-ratio**	***p***
Intercept, γ_{00}	81.63	3.35	24.36	<.001
District, γ_{01}	-1.56	.50	-3.14	<.01
TSELI, γ_{02}	.04	.02	1.88	<.05
Ethnicity, γ_{10}	-2.89	.45	-6.38	<.001
ELL, γ_{20}	-4.44	.75	-5.91	<.001
SPED, γ_{30}	-10.0	.58	-17.08	<.001
Lunch, γ_{40}	-4.65	.48	-9.64	<.001
Random Effects	**Variance components**	**SD**	**$\chi 2$**	***p***
Intercept, τ_{00}	21.35	4.62	917.74	<.001
Residual, σ^2	121.90	11.04		
Model Fit	**Value**	**Parameters**	**Variance Accounted for**	
Deviance	29369.37	2	56%	

The regression coefficients for all variables retained in the final hierarchical model were all significant, evidencing their utility in predicting student reading achievement. The final model accounted for 56% of the variation in student reading achievement.

The primary focus of this research question was to determine the impact of teacher efficacy for literacy instruction on reading achievement. While TSELI scores did not account for as much variation in reading achievement (2%) as some of the more classical predictors of student learning outcomes (in the study sample, ethnicity alone accounted for 22% of the variation in reading achievement; lunch status 19%; ELL status 24%, and SPED status 26%) and its effect was relatively small (a one-point gain in TSELI scores was associated with a .04 increase in reading achievement in the final model), its retention in the final model as a significant predictor evidences the importance of this construct.

Summary of Findings

This study investigated the relationship between various teacher and classroom characteristics and teachers' sense of efficacy for literacy instruction as well as the impact of efficacy for literacy instruction on student reading achievement. Teacher efficacy for literacy instruction did not significantly differ across grade levels; average TSELI scores remained stable between grades 3 through 8. While differences existed between districts participating in the study, within districts, significant differences were not found across grade level. Teacher characteristics found to influence efficacy for literacy instruction included the length of a teacher's tenure in the profession and the highest level of education completed. Educators teaching for more than 10 years and those with graduate degrees had significantly higher levels of efficacy for literacy instruction. Classroom characteristics also were evaluated to ascertain how student demographic variables aggregated across the classroom influenced teacher efficacy for literacy instruction. Classroom teachers with higher numbers of Caucasian students and fewer students receiving either free or reduced school lunches had significantly higher levels of efficacy for literacy instruction. Finally, the impact of efficacy for literacy instruction on student reading achievement was analyzed. As a preliminary step, it was determined that the students of teachers with higher levels of subject-specific efficacy scored significantly higher, on average, on a standardized measure of reading achievement. However, the effect was small in that efficacy for literacy instruction accounted for 1% of the variation in reading achievement. Finally, a hierarchical analysis was performed to evaluate the predictive power of efficacy for literacy instruction. While the classical predictors of student achievement—variables such as ethnicity, SES, and SPED status—have a greater impact on reading scores than teacher perceptions, efficacy for literacy instruction was a significant predictor of reading achievement in the hierarchical model that best fit the study data.

CHAPTER FIVE

Findings, Conclusions, and Implications

Background

The importance of teaching students to read and to comprehend what they read is critical, as every child must develop this skill to be successful in school. Regardless of the subject matter taught, children must be able to read and comprehend what they are reading. However, far too many American students struggle to read at a level that will allow them to excel in school and in their future workplaces. In fact, the most recent National Assessment of Educational Progress (2011) revealed that a quarter of eighth graders read at “below basic” levels. These students have little chance of succeeding in the demanding courses of high school without interventions that will considerably increase their ability to comprehend complex material, increase their vocabularies, and develop strong study skills.

A number of variables have been found to influence student achievement, but numerous studies have documented that teacher quality, of all of the variables that can be controlled, has the greatest effect on student learning (Goldhaber, 2002; Sanders & Rivers, 1997; Sanders, Saxton, & Horn, 1997). Through the years, one construct assisting in the measurement of teacher quality has been the construct of teacher efficacy. Teacher efficacy has been linked to a number of positive outcomes, including increased student achievement. Effectively measuring the construct of efficacy, however, has been challenging at best for researchers and often times elusive, as the findings from scientific inquiries involving efficacy have often been as varied as the studies themselves. In recent years, there has been increased attention on the measurement of domain-specific efficacy instruments. Bandura (2001) offered that, “Scales of perceived self-efficacy must be tailored to the particular domains of functioning that are the object of interest”

(p. 10). In response to this, Tschannen-Moran and Johnson (2011) developed the Teacher Sense of Efficacy for Literacy Instruction (TSELI) scale to measure teachers' sense of efficacy for literacy instruction. Given the documented power of the construct of teacher efficacy and the importance of reading as it relates to learning outcomes, this study sought to examine the relationship between the efficacy of teachers and student reading achievement.

Purpose of the Study

The purpose of this study was to examine the relationship between various teacher and classroom variables on subject-specific efficacy and to investigate the impact of teachers' efficacy beliefs for literacy instruction on reading achievement for third- through eighth-grade students. It was hypothesized that high levels of efficacy for literacy instruction were associated with increased student learning outcomes in reading in grades three through eight. Additionally, contextual variables, including the number of years teaching, highest level of education, gender, and professional-development experiences of participating teachers were collected in order to ascertain their influence on efficacy. Differences in classroom variables based on student demographic information such as gender, free-and-reduced lunch status, special-education status, ethnicity, and English language learner status also were evaluated with regard to their impact on teacher efficacy for literacy instruction.

Research Questions

1. Does a teacher's level of sense of efficacy for literacy instruction decrease as grade levels increase?
2. How do teacher characteristics such as gender, the number of years teaching, education level, and professional development influence teachers' sense of efficacy for literacy instruction?
3. How do classroom-specific characteristics, such as gender, English language learner status, special-education status, free-and-reduced lunch status, and students' ethnicity influence teachers' sense of efficacy for literacy instruction?
4. How does a teacher's sense of efficacy for literacy instruction impact student reading achievement?

Discussion of Findings

This study examined four research questions pertaining to the Teacher Sense of Efficacy for Literacy Instruction Scale (TSELI). The TSELI is a relatively new measure, and extensive research-based evidence documenting its validity or its relationship to student outcomes does not exist. Therefore, employing the TSELI for this study provided important information for research based around this measurement tool. This study's first three research questions sought to uncover what variables influenced teachers' sense of efficacy for literacy instruction. To what extent did efficacy for literacy instruction differ across various levels of teacher characteristics and classroom contextual variables, based on aggregated student demographic information? The focus shifted for the fourth study question to assess the impact of efficacy for literacy instruction on student achievement in reading. Specifically this question centered on the extent to which teacher efficacy for literacy instruction predicted student reading outcomes.

Research question one addressed whether efficacy for literacy instruction varied across grade levels, specifically hypothesizing that levels of efficacy would decrease as grade levels increased. This hypothesis was based on previous research examining efficacy using general (as opposed to domain-specific) measures that found lower grade-level teachers had higher levels of efficacy. As detailed in Chapter Four, the null hypothesis of no difference between grade-level teachers' sense of efficacy for literacy instruction was unable to be rejected, as the grade-level TSELI scores were not found to be significantly different from each other.

Examining teachers' levels of self-efficacy for literacy instruction across grade levels was included as a research question for several reasons. First, Tschannen-Moran & Johnson (2011) found school level to exhibit a significant difference regarding efficacy for literacy instruction, with elementary teachers demonstrating higher levels of the construct than those in middle and high schools. Furthermore, Tschannen-Moran & McMaster (2010) found that primary grade teachers tended to view themselves as doing a good job at teaching in general, specifically teaching reading. Teachers in the Tschannen-Moran & McMaster study showed high mean scores of both teacher self-efficacy in general and teacher self-efficacy for reading instruction. Typically, primary-grade teachers felt more comfortable teaching reading, because it was so central to the task of teaching primary grades.

This study found that teachers' level of efficacy for literacy instruction was remarkably consistent across grade levels, with TSELI scores ranging from 146.0 (grade 5) to 152.7 (grade 7), with an overall sample mean of 149.1. This result indicated that for this sample, efficacy for literacy instruction was independent of grade level. That is, teacher perceptions regarding their ability to effectively teach literacy did not vary as a function of grade level. This result seemed to contradict the finding regarding grade and level of efficacy from the only published study evaluating the TSELI (Tschannen-Moran & Johnson, 2011). It should be noted that the Tschannen-Moran & Johnson (2011) study compared TSELI scores for samples of elementary- and middle-school teachers, while the current study examined teacher efficacy at each grade level.

Teacher demographic and experience variables, including number of years teaching, education level, gender, and professional-development experiences were the focus of research question two. Of the demographic and experience variables evaluated, education level and number of years teaching were found to be the strongest factors influencing efficacy for literacy instruction, with longer-tenured teachers and those with graduate degrees exhibiting higher levels of self-efficacy. Scores on the TSELI did not differ significantly for males and females.

The findings for all three of these demographic teacher variables contradicted the findings of Tschannen-Moran and Johnson (2011), who found that both the number years taught and education level were unrelated to efficacy for literacy instruction beliefs, while males and females did exhibit significant differences. The findings from the current study also differed from previous work by Bandura (1997), which maintained that efficacy beliefs tended to be set early and remained stable and resistant to change without some kind of shock to provoke reassessment. However, the work of Bandura (1997) was based on examinations of general measures of efficacy, as opposed to domain-specific measures, such as the TSELI. Given the specific nature and highly-specialized focus of the TSELI, it seems plausible that either higher levels of formal educational training or the sum of experiences and knowledge gained through many years teaching could be the shock-provoking reassessment of efficacy beliefs to which Bandura referred. Further research is needed to examine the relationship of these two variables, given the conflicting results with the Tschannen-Moran and Johnson (2011) study.

Research question two also considered professional- and personal-development experiences of teachers. It should be reiterated that the sample size for these analyses was considerably smaller than the sample size for the other analyses supporting the second research questions. Questions pertaining to the professional- and personal-development experiences of teachers were asked as follow-up questions to the TSELI and demographic information, and not all teachers who originally participated in the survey answered the follow-up questions. The results of the follow-up questions relating to professional- and personal-development experiences were quite surprising. Both the perception of the quality of undergraduate teacher-preparation programs and perceptions of the quality of district professional-development sessions had an inverse relationship with efficacy scores: those who rated their university and district professional-development experiences lowest scored highest on the TSELI, although neither was found to be statistically significant. Lack of significance was likely due to the small sample size contained within each cell of the analyses.

These findings call into question the impact that knowledge base has on this literacy specific measure. The TSELI (Appendix D) asks very specific questions pertaining to literacy instruction, and it is well documented both anecdotally and through research-based methods that effectively teaching literacy is a challenging task and a complicated process. These results bring into question whether the more one knows about literacy instruction, the more accurately that person can complete the TSELI. In other words, do these teachers know what they don't know? Teachers who indicated having positive views of their undergraduate training and professional-development experiences actually may be able to respond more critically to the TSELI because they know the areas in which they are challenged when it comes to literacy instruction. The findings regarding professional and personal development, while not statistically significant due to small sample size, were curious and deserve a closer examination in future research.

Moving from teacher variables to student characteristics, research question three focused on how classroom-contextual variables affected efficacy for literacy instruction. Student-level variables examined included gender, ethnicity, SES (as measured by the proxy of lunch-program status), special-education status, and English language learner status. Of these variables, free-and-reduced lunch status and ethnicity showed statistically significant differences on TSELI scores. Results of this study found that teachers' sense of efficacy for literacy instruction was higher when the socioeconomic status of his or her classroom was higher. Additionally, TSELI scores were higher for teachers who had the highest percentages of Caucasian students in their classes.

The findings that SES and ethnicity negatively impacted TSELI were disheartening and demand additional research. While both ethnicity and SES have been powerful predictors of achievement and have been the focus of much educational research through the years, prior research has been mixed when it comes to the effect of contextual variables on teacher efficacy, and contextual factors have been posited to have a minimal impact, especially as teachers gain more experience. This study found the contextual variables of student SES and ethnicity to have among the strongest exhibited effects on efficacy for literacy instruction of any variables analyzed.

While discussing student-level characteristics and teacher efficacy, it is important to highlight the distinction in the research between individual teacher efficacy and collective teacher efficacy. This study utilized a measure of individual efficacy, so emphasis was not placed on collective efficacy. However, the collective efficacy, the perception of teachers in a school that the faculty as a whole can have a positive effect on students (Goddard, Hoy & Woolfolk-Hoy, 2000), is an important construct in emerging research in education because of its positive association with student achievement. Bandura (1993) reached two important conclusions in his seminal research: (a) student achievement is significantly and positively related to collective efficacy, and (b) collective teacher efficacy has a greater effect on student achievement than does student socio-economic status. Goddard and Woolfolk-Hoy (2000) confirmed Bandura's findings, reporting that collective teacher efficacy was more predictive of elementary students' math and reading achievement than was gender, ethnicity, and socio-economic status. Such findings have not been supported through the use of an individual teacher efficacy measure. This distinction calls one to consider how collective and individual efficacy differs in both construct and measurement.

Having engaged in a thorough treatment of factors influencing efficacy levels, the focus of research question four shifted from what variables affect on teachers' sense of efficacy for literacy instruction to the impact of a teacher's TSELI score have on student achievement. This question was of interest, because studies conducted using general efficacy measures (Anderson, Green and Loewen, 1988; Tranz & Gibson, 1986) found significant relationships between teacher efficacy and reading achievement. However, this is the first study to examine the relationship between the literacy-specific measure of efficacy and student reading achievement. Tschannen-Moran and Johnson (2011) conducted an analysis of teacher variables on the TSELI but did not study its impact on student achievement in reading.

Two analytic approaches were implemented to address this study question. The first approach utilized ANOVA to determine whether student reading achievement differed significantly for teachers with low, mid, and high levels of efficacy for literacy instruction.

Results of this analysis showed that differences in reading achievement were significant for various levels of efficacy. The second component of this analysis sought to address whether efficacy for literacy had a meaningful impact on reading achievement. In other words, the independent contribution of efficacy for literacy instruction on student reading achievement was examined through use of hierarchical linear modeling. The HLM analysis found efficacy for literacy instruction to be a significant predictor, although the amount of variance in student reading achievement that the construct accounted for was small. Interpretation of the regression coefficient for the TSELI variable in the final retained HLM model found that a one-point change in TSELI scores is associated with a .04 point increase in student reading achievement. The findings of this analysis show that there was a statistically significant relationship between teachers' sense of efficacy for literacy instruction and student achievement gains in reading.

Limitations of the Study

Several methodological and theoretical limitations must be addressed regarding this study.

1. The study sample consisted of teachers from large districts within the state of Kansas who were identified by the researcher. While drawing participants from large districts was designed to minimize the differences that may have existed between urban/suburban and rural districts, the study sample may not be representative of the general population, thereby limiting researchers' ability to generalize results.
2. Given the voluntary nature of the data gathering method, it may be that the teachers who responded differed in some systematic way related to the variables under study than those who did not respond.
3. The study did not control for prior differences in student reading ability, so exhibited differences between levels of independent variables may have been impacted by pre-

existing differences among groups. While this potential issue was not addressed statistically, efforts were made to obtain the largest sample possible for the current study. While both the teacher and student samples were sufficient for the analytic procedures employed, some techniques involved the analysis of multiple factors with multiple levels, which resulted in imbalanced cell sizes. To combat any potential issues, data were checked thoroughly for violations of assumptions for performing statistical tests prior to all analyses.

4. With regard to the measurement of teacher efficacy, the TSELI is a new tool designed specifically to measure efficacy for literacy instruction. While reliability coefficients and validity evidence have been documented in a single published study (Tschannen-Moran & Johnson, 2011), the survey is not supported by a significant body of applied research. Furthermore, the TSELI relied on self-reported responses. Previous research found that teachers often overestimated their content knowledge and ability, which could affect the validity of the measurement of this construct (Cunningham et. al., 2004; Tschannen-Moran & Johnson, 2011).
5. The distribution of teacher self-reported scores on the TSELI was centered in the higher end of the score scale, with the average item response being approximately 7 on a scale of 1 to 9. While it is entirely possible that the sample consisted of highly efficacious teachers, the elevated levels of responses call into question the ability of the TSELI to truly discriminate—in the statistical sense—between teachers with varying levels of efficacy.

Implications

This study has several important implications, the most important being its contribution to the literature pertaining to the relationship between teacher efficacy for literacy instruction and student achievement. This study utilized the TSELI, and to date, no studies have investigated the relationship between the TSELI and student achievement in reading. Interestingly, a number of the teacher-level findings of this study contradicted with those of the Tschannen-Moran and Johnson (2011) study. These contradictions highlight the need for further examination. It would be beneficial for the TSELI to undergo additional reliability and validity testing. Furthermore, additional studies employing the TSELI would help to broaden the literature base.

Measuring the construct of efficacy has been problematic since its inception. Hanson (2001) notes, "The study of teacher efficacy has suffered from poor construct validity issues" (p. 17). One measurement issue that may have affected the results of this study was mentioned above, the fact that the TSELI is a new measure and would benefit from undergoing additional reliability and validity testing. Additionally, the self-reporting nature of most efficacy measures creates an underlying problem in the research, including this study. When teachers are asked to rate their own abilities, they may overestimate their abilities. The distribution of teacher self-reported scores on the TSELI was centered in the higher end of the score scale, with the average item response being approximately 7 on a scale of 1 to 9. While it is entirely possible that the sample consisted of highly efficacious teachers, the elevated levels of responses call into question the ability of the TSELI to truly discriminate—in the statistical sense—between teachers with varying levels of efficacy. Does the TSELI engender responses that allow teachers overestimate their abilities? Would it be more effective to shift the focus of questions from: 'To what extent can you do *this*?' to 'How often do you do *this*?'? It is interesting to note that the

Tschannen-Moran and Johnson (2011) study produced similar results in terms of average item ratings.

Furthermore, findings in the literature have not been consistent, as a number of studies have been cited that uncover no significance between teacher efficacy and achievement (Ashton & Webb, 1986; Borton, 1991). Lack of consistency of methodological approaches used across studies has created problems in obtaining accurate results. All of these problems combined illuminate why efficacy has been a difficult construct to measure and to analyze precisely.

While this study found that reading levels were higher for students of teachers with higher levels of efficacy, and efficacy was retained in the final hierarchical model for predicting reading achievement, the individual contribution of efficacy for literacy instruction was practically insignificant. Considering this, one must ask if the effort required to increase efficacy is energy best spent. Or, is energy better expended addressing variables that have been found to have a larger impact on achievement? Teacher efficacy is undoubtedly an interesting construct. However, it appears that additional research must be conducted to better understand the construct.

In 1992, Raudenbush, Rowen and Cheong challenged teacher-efficacy research and embarked on a study to examine the casual relationship between teachers' perceived self-efficacy and student achievement, noting, "positive feelings of self-efficacy are necessary (but not sufficient) for effective teaching" (p. 151). Bandura (1986) explained that, "some situations require greater skill and more arduous performances, or carry greater risk of negative consequences, than others" (p. 311), and teacher efficacy can vary from situation to situation. Raudenbush, Rowen, and Cheong found that teachers were more efficacious when teaching honors classes, less efficacious in teaching academic courses, and least efficacious in vocational and general track courses. In the Raudenbush, Rowen, and Cheong study, 44% of the variance in teachers' perceived self-efficacy was within-teacher variance, which included which tracks and subject specialization they taught.

The results of the current study echo the findings of the study by Raudenbush and his colleagues (1992), which found teacher efficacy to be influenced significantly by variables that have been found to be negatively related to achievement. This finding begs consideration of whether self-reported teacher efficacy is a revelation of one's own perceived abilities or reflects the abilities of one's students.

Recommendations for Future Research

This study's findings have implications that may prove interesting for future research. Some suggestions for further research include:

1. Use the Teacher Sense of Efficacy Scale (TSES), which measures teachers' efficacy beliefs for more general aspects of teaching, and the Teacher Sense of Efficacy for Literacy Instruction Scale (TSELI) to conduct analysis similar to those employed in this study to both scales. It would be beneficial to see how the teachers scored on the general measure versus the literacy specific measure.
2. It would be interesting to see if the results change if a study used a larger sample involving more school districts. The current study used mainly large, urban/suburban districts. It would be beneficial to include additional districts of similar make-up and smaller, rural districts as well.
3. Given the emerging research base around collective efficacy, it would be beneficial to develop a collective efficacy measure specific to literacy. Once the measure was validated, a study could examine how collective efficacy for literacy instruction

affected literacy achievement. It would be useful to compare how results from the collective efficacy scale differed from the individual efficacy scale.

4. A similar study could be conducted using a measure other than the Kansas Assessment to measure student achievement in reading or a measure in addition to the Kansas Assessment.
5. It would be very beneficial to conduct a mixed-methods study pertaining to teachers' sense of efficacy for literacy instruction. Teacher-efficacy research, including this study, has been dominated by quantitative methodologies. Given the self-reporting of the TSELI (and most efficacy measures), it would be powerful to use fewer teachers but to include classroom observations, interviews, and other qualitative methods paired with the TSELI or a different efficacy scale. Then the teacher's scale could be compared with the researcher's rating based on his/her experiences gleaned through the qualitative analysis..
6. Rather than study the TSELI and its effect on student achievement, the TSELI could be used as a basis for professional development. Teachers could take the TSELI in a pre-test form, receive targeted professional development, and take a post test to see if their TSELI scores changed.
7. Use the TSELI to survey classroom teachers, special-education teachers, English language teachers, reading coaches, and interventionists.

Conclusion

Determining ways to increase student reading ability continues to be an important charge for those involved in education. The construct of teacher efficacy has been viewed as a variable that plays a significant role in student achievement. Tschannen-Moran and Woolfolk Hoy (2001) defined teacher efficacy as a teacher's "judgment of his or her capabilities to bring about desired outcomes of student engagement and learning, even among students who may be difficult or unmotivated" (p. 283). The current study examined factors influencing teacher efficacy for literacy instruction and evaluated the impact of the construct on student achievement. Domain-specific efficacy measurement is new to literacy instruction, and this study attempted to contribute to the research literature around this topic. While the practical significance of efficacy for literacy instruction on reading achievement was not compelling, the findings are an important contribution to the body of literature on how teachers affect student achievement. However, because of the relatively small magnitude of the relationships found in the current study, the examination of teacher efficacy for literacy instruction and the influence of those beliefs on student achievement is an area that requires further analysis. As more researchers use the findings and recommendations from the current study to inform new investigations of the complex relationship between teacher efficacy and student achievement in reading, teacher educators, policymakers, teachers, and administrators will be better informed as they continue to work towards improving reading-achievement scores.

Teacher efficacy has been used to measure teacher quality, and according to the literature (Ashton & Webb, 1884; Bandura, 1997, 2001; Gibson & Dembo, 1984; Tschannen-Moran, Hoy, & Hoy, 1998; Tschannen-Moran, & Hoy, 2001), has been found to correlate with student achievement. The purpose of this study was to investigate the relationship between third- through eighth-grade classroom teachers' sense of efficacy for literacy instruction and student reading achievement. This study found that although certain teacher- and student-level variables did have a statistical significance on TSELI, their overall impact was small. Additionally, as scores on the

TSELI increased, so did reading achievement. Results of this study contribute to the research examining the TSELI and student achievement.

REFERENCES

Allinder, R.M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education, 17,* 86-95.

Anderson, R., Green, M., & Loewen, P. (1988). Relationship among teachers' and students' thinking skills, sense of efficacy, and student achievement. *Alberta Journal of Educational Research, 34(2),* 148-165.

Armor, D., Conroy-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E. & Zellman, G. (1976). *Analysis of the school preferred reading programs in selected LosAngeles minority schools* (Rep. No. R-2007-LAUSD). Santa Monica, CA: RAND. (ERIC Document Reproduction Service No. 130 243).

Ashton, P.T., Olejnik, S., Crocker, L., & McAuliffe, M. (1982). *Measurement problems in the study of teachers' sense of efficacy.* Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Ashton, P.T., Buhr, D., & Crocker, L. (1984). Teachers' sense of efficacy: a self-or-norm referenced construct? *Florida Journal of Educational Research, 21(1),* 29-41.

Ashton, P.T., & Webb, R.B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement.* New York: Longman.

Bandura, A. (1986). *Social foundation of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191-215.

Bandura, A. (1997). *Self- efficacy: The exercise of control.* New York: W.H. Freeman and Company.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52,* 1-26.

Basken, P. (2006) States have more schools falling behind. *Bloomberg News,* 1. Retrieved from: http://www.utsystem.edu/news/clips/dailyclips/2006/0326-0401/K12-BN-More-032906.pdf

Baumann, J., Hoffman, J., Duffy-Hester, A., & Moon, J. (2000). Teaching teachers to teach reading; paradigm shifts, persistent problems and challenges. In Kamil, P. Mosenthal, P. Pearson & R. Barr (Eds.), *Handbook of reading research, Vol. 3* (pp. 719-742). Mahwah, NJ: Erlbaum.

Berman, P., McLaughlin, M., Bass, G., Pauly, E., & Zellman, G. (1977). *Federal programs supporting educational change: Vol. VII. Factors affecting implementation and continuation.* (Rep. No. R-1589/7-HEW). Santa Monica, CA: RAND. (ERIC Document Reproduction Service No. 140 432).

Biancarosa, G., & Snow, C. E. (2004). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education. Retrieved May 24, 2006, from: http://www.ncte.org/edpolicy/literacy/research/122355.htm

Borton, W.M. (1991). *Empowering teachers and students in a restructuring school: A teacher efficacy interaction model and the effect on reading outcomes.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. (ERIC Reproduction Document Service No. ED 335341).

Burley, W.W., Hall, B.W., Villeme, M.G., & Brockheimer, L.L.(1991). *A path analysis of the mediating role of efficacy in first- year teachers' experiences, reactions and plans.* Paper presented at the annual meeting fo the American Educational Research Association, Chicago, IL.

Cantrell, S.C., & Calloway, P. (2008). *High and low implementers of content instruction: Portraits of teacher efficacy, 24,* 2739-1750.

Cohen, J. (1988). *Statistical power analysis for behavior sciences.* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Coladaraci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education, 60,* 323-337.

Common Core State Standards Initiative. (2010). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. Developmental Psychology, 33, 934-945.

Cunningham, A.E., Perry, K.E., Stanovich, K.E., & Stanovich, P.J. (2004). Disciplinary knowledge of K-3 teachers and their knowledge calibration in the domain of early literacy. *Annals of Dyslexia, 54,* 139-169.

Czerniak, C.M., & Shriver, M.L. (1994). An examination of preservice teachers' beliefs and behaviors as related to self-efficacy. *Journal of Science Teacher Education, 5*(3), 77-86.

DaCosta, J.L., & Riordan, G. (1996). *Teacher efficacy and the capability to trust.* Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

DeMusquita, P.B., & Drake, J.C. (1994). Educational reform and the self-efficacy beliefs of teachers implementing nongraded primary school programs. *Teaching and Teacher Education, 10,* 291-302.

Darling-Hammond, L. (1997). Toward what end? The evaluation of student learning for the improvement of teaching. In J. Millman (Ed.), *Grading teachers, grading schools. Is student achievement a valid evaluation measure?* (pp. 248-263). Thousand Oaks, CA: Corwin.

Darling-Hammond, L. (2000). Teacher quality and student achievement. A review of state policy evidence. *Educational Policy Analysis Archives, 8*(1). Retrieved from http://epaa.asu.edu/epaa/v8n1

Egyed, C., & Short, R. (2006). Teacher self-efficacy, burnout, experience and decision to refer a disruptive student. *School Psychology International, 27(4)*, 462-474.

Elementary and Secondary Education Act of 1965. RL. No. 89-10 (1965). *Elementary and secondary education* . (2010, December 16). Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/index.html

Emmer, E. (1990). *A scale for measuring teacher efficacy in classroom management and discipline.* Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.

Enochs, L.G., Scharman, L.C., & Riggs, I.M. (1995). The relationship of pupil control to preservice elementary science teacher self-efficacy outcome expectancy. *Science Education, 79(1),* 63-75.

Evans, E.D., & Tribble, M. (1986). Perceived teaching problems, self-efficacy and commitment to teaching among preservice teachers. *Journal of Educational Research, 80(2),* 81-85.

Fritz, J. J., Miler-Heyl, J., Kreutzer, J. C., & MacPhee, D. (1995). Fostering personal teaching efficacy through staff development and classroom activities. *The Journal of Educational Research, 88*(4), 200-208.

Gibson, S., & Dembo, M. (1984). Teacher efficacy: a construct validation. *Journal of Educational Psychology, 76,* 569-582.

Graham, S., Harris, K.R., Fink, B., & MacArthur, C.A. (2010). Teacher efficacy in writing: A construct validation with primary teachers. *Scientific Studies in Reading, 5*(2), 177-202.

Goddard, R.D., Hoy, W.K., & Woolfolk, A. (2000). Collective teacher efficacy: Its meaning, measure, and effect on student achievement. *American Education Research Journal, 37*(2), 479-507.

Goldhaber, D. (2002.) The mystery of good teaching: Surveying the evidence on student achievement and teachers' characteristics. *Education Next* 2(1): 50–55.

Guskey, T.R. (1981). Measurement of the responsibility teachers assume for academic successes and failures in the classroom. *Journal of Teacher Education, 32(3),* 44-51.

Guskey, T.R. (1986). Staff development and the process of teacher change. *Educational Researcher, 15,* 5-12.

Guskey, T.R. (1987). Context variables that affect measures of teacher efficacy. *Journal of Educational Research, 81(1),* 41-47.

Guskey, T.R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 4,* 63-69.

Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading*, 3(3), 231-257.

Hansen, (2001). Teacher self-efficacy: Substantive implications and measurement dilemmas. Paper presented at the annual meeting of the Educational Research Exchange, Texas A&M University, College Station, Texas.

Hansen, S. (2006). I can't do everything! English teachers' efficacy beliefs. *International Journal of Learning. 13*(7), 53-60.

Hanushek, E.A., Kain, J.F., & Rivkin, S.G. (1999), "Do Higher Salaries Buy Better Teachers?" Working Paper. Retrieved from: http://www3.amherst.edu/~sgrivkin/papers/salpap.pdf

International Reading Association (2004). *Standards for reading professionals.* Newark, DE: International Reading Association.

Institute for Student Achievement (2010). Transforming America's high schools Retrieved from: http://www.studentachievement.org/WhyISA.aspx?id=fqhYz7CKTzI=

Jerald, C. D. (2007, January). *Believing and achieving.* Washington, DC: The Center for Comprehensive School Reform and Improvement.

Johnson, D. (2010). Learning to teach: the influence of a university-school partnership project on pre-service elementary teacher's self-efficacy for literacy instruction. *Reading Horizons, 50*(1), 23-48.

Kansas State Department of Education (2008). *Assessments.* Retrieved September 10, 2011, from http://www.ksde.org/Default.aspx?tabid=420

Maloch, B., Flint, A., Eldridge, D., Harmon, J., Loven, R., Fine, J., et al. (2003). Understanding, beliefs, and reported decision making of first-year teachers from different reading teacher preparation programs. *The Elementary School Journal, 103*(5), 431-459.

Meijer, C., & Foster, S. (1988). The effect of teacher self-efficacy on referral chance. *Journal of Special Education, 30,* 152-186.

Midgley, C., Feldlaufer, H., & Eccles, J. (1989). Change in teacher efficacy and student self-related-beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology, 81,* 247-258.

Mone, M.A., Baker, D.D., & Jeffries, F. (1995). Predictive validity and time dependency of self-efficacy, self-esteem, personal goals and academic performance. *Educational and Psychological Measurement, 55(5),* 716-727.

Moore, D.W., Bean, T.W., Birdyshaw, D., & Rycik, J. A. (1999). Adolescent Literacy: A Position Statement for the Commission on Adolescent Literacy of the International Reading Association. Newark, DE: International Reading Association.

Moore, William P., & Esselman, Mary E. (1992). Teacher efficacy, empowerment, and a focused instructional climate: Does student achievement benefit? Paper presented at the annual meeting of the Annual Conference of the American Educational Research Association: San Francisco, CA. (ERIC Document Reproduction Service No. ED350252).

Nathanson, S., Pruslow, J., & Levitt, R. (2008). The reading habits and literacy attitudes of in-service and prospective teachers. *Journal of Teacher Education, 59,* 313-321.

National Assessment of Educational Progress (2011). *Nation's report card findings in brief reading and mathematics.* Retrieved from: http://nces.ed.gov/whatsnew/

National Council of Teachers of English and the International Reading Association, (1996). *Standards for the English language arts.* Urbana, IL: NCTE.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence based assessment of scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington DC: U.S. Government Printing Office.

National Commission on Excellence in Education. (1983). *A Nation at Risk. The Imperative for Education Reform.* U.S. Government Printing Office (*http://www.ed.gov/pubs/NatAtRisk/risk.html*).

No Child Left Behind Act of 2001. (2001). Public Law107-110.

Pajares, F., & Miller, M.D. (1994). Role of self efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86(2),* 192-203.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66(4),* 543-578.

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Continuing Educational Psychology, 24,* 124-139.

Peterson, D.R. (2003). Unintended consequences: Ventures and misadventures in the education of professional psychologists. *American Psychologist*, *58*(10), 791-800.

Pintrich, P.R., & Schunk, D.H. (1996). *Motivation in education: Theory, research and application.* Englewood Cliffs, NJ: Merrill.

Poggio, A.J., Irwin, P.M, Yang, X., Glasnapp, D.R., & Poggio, J.P. (2007). Technical Manual for the Kansas Assessments in Reading and Mathematics. Kansas State Department of Education.

Race to the top. (2011, November 3). Retrieved from http://www2.ed.gov/programs/racetothetop/index.html

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading.* Santa Monica, CA and Washington, DC: RAND Corporation.

Rasinski, T. (2003). *The fluent reader: Oral reading strategies for building word recognition*. New York, NY: Scholastic Professional.

Raudenbush, S., Rowan, B., & Cheong, Y.F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education, 65,* 150-167.

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

Reeves, D. (2004). Accountability for learning: How teachers and school leaders can take charge. Alexandria, VA: ASCD.

Rigg, I.M., & Enochs, L.G. (1990). Toward the development of an elementary teacher's science teaching efficacy beliefs instrument. *Science Education, 74(6),* 625-637.

Rose, J.S., & Medway, F.J. (1981). Measurement of teachers' beliefs in their control over student outcomes. *Journal of Educational Research, 74.* 185-190.

Ross, J.A. (1998). The antecedents and consequences of teacher efficacy. In J. Brophy (Ed.), *Research on teaching* (Vol. 7, pp. 49-74). Greenwich, CT: JAI.

Ross, J., Cousins, B., & Gaddalla, T. (1996). Within-teacher predictors of teacher efficacy. *Teacher and Teacher Education*, *12*, 385-400.

Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80,* 1-28.

Sanders, W.L., & Rivers, J.C. (1997). Cumulative and residual effects of teachers on future student academic achievement. Research Progress Report. Knoxville: University of Tennessee Value-Added Research and Assessment Center.

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

Schunk, D.H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J.S. Eccles (Eds.), *Development of achievement motivation* (pp. 15-31). New York: Academic Press.

Shanahan, T. (2011, October 18). *Common core state standards English Language Arts.* Retrieved from http://www.shanahanonliteracy.com/

Shaughnessy, J. (2004). Is your school ready to tackle the emotional issues around change? *Emotional Literacy Update, 7*, 8-9.

Snijders, T., & Bosker, R. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational Statistics*, *18*(3), 237-59.

Snow, C. E., Burns, S. M., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Soodak, L.C., & Podell, D.M. (1993). Teacher efficacy and student problems as factors in special education referral. *The Journal of Special Education, 27(1),* 66-81.

Striving readers. (2010, March 8). Retrieved from http://www2.ed.gov/programs/strivingreaders/index.html

Szabo, S., & Mokhtari, K. (2004). Developing a reading teacher efficacy instrument for teacher candidates: A validation study. *Action in Teacher Education, 26*(3), 59-72.

Timperley, H.S., & Phillips, G. (2003). Changing and sustaining teachers' expectations through professional development in literacy. *Teaching and Teacher Education, (19),* 627-641.

Tournaki, N., & Podell, D.M. (2005). The impact of student characteristics and teacher efficacy on teachers' predictions of student success. *Teaching and Teacher Education,* *21*(3). 299-314.

Tranz, S.M., & Gibson, S. (1986). *Effects of efficacy on academic achievement.* Paper presented at the annual meeting of the California Educational Research Association: Marina Del Ray, CA. (ERIC Document Reproduction Service No. ED281853).

Tschannen-Moran, M., Woolfok Hoy, A., & Hoy, W.K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202-248.

Tschannen-Moran, M., & Woolfok Hoy, A. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and teacher education, 17,* 783-805.

Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy.

Tschannen-Moran, M., & Johnson, D. (2011). Exploring literacy teachers' self-efficacy beliefs: Potential sources at play. *Teaching and Teacher Education, 27*(4). 751-761.

Tschannen-Moran, M., & MacFarlane, B. (2011). Teacher self-efficacy in the English language arts classroom. In Lapp, D., & Fisher, D., (Eds), *Handbook of research on the teaching of the English language arts.* 3rd. ed. New York: Routledge.

U.S. Congress. *National Defense Education Act of 1953,* P.L. 85-864. 85th Congress, September 2, 1958. Washington, DC: GPO.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71,* 3-25.

Weiner, B. (1992). *Human motivation: Metaphors, theories and research.* Thousand Oaks, CA: Sage.

Woolfolk, A.E., & Hoy, W.K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology. 82*,81-91

Yeung, K.W., & Watkins, D. (2000). Hong Kong student teachers' personal construction of teaching efficacy. *Educational Psychology, 20*(2), 213-235.

APPENDIX A

LETTER TO TEACHER PARTICIPANTS

Letter to Teacher Participants

Dear Teacher Participant:

You have been invited to participate in a study about the impact of teacher efficacy on student achievement in reading. This study is being conducted by Jennifer Poggio, a doctoral candidate at the University of Kansas, as part of her doctoral dissertation. Her faculty sponsor is Dr. Arlene Barry, who can be contacted at:

KU School of Education
208 Joseph R. Pearson Hall
1122 West Campus Rd.
Lawrence, KS 66045-3101
785 864 9661
abarry@ku.edu

You have been selected for this study because you are a 3rd- 8th grade teacher or are a reading specialist in your district. The purpose of this study is to examine the relationships between teachers' beliefs about their influence on students' ability to affect student learning and Kansas Assessment data in the area of reading.

If you agree to participate in this study, you will be asked to complete the following survey. Completion of the survey indicates your willingness to participate in this project and that you are over the age of 18.

Participation in this study will involve approximately 15 minutes of your time. There are no consequences for choosing not to participate or for withdrawing from the study. This research may help us better understand how a teacher's sense of efficacy affects student achievement in reading.

Confidentiality of your identity will be strictly maintained. The information from this study will only be accessible by the researcher and your district assessment coordinator. All teacher and student information will be coded so that no individual can be identified by the researcher. Similarly, results from this study will be released in a way that no individual answers can be identified. It is possible, however, with internet communications, that through intent or accident, someone other than the intended recipient may see your response. All information gathered will be kept secure, under physical lock and key, for three years after the research has been completed. After three years, it will be destroyed with a paper shredder and recycled.

For any questions regarding this study or problems with your role as a teacher participant, you may contact Jennifer Poggio or Dr. Arlene Barry, at the address above. If you should have any additional questions about your rights as a research participant, you may call (785) 864-7429 or email mdenning@ku.edu, the Human Subjects Committee Lawrence Campus (HSCL).

Sincerely,

Jennifer Poggio

jenniferpoggio@ku.edu

(913) 322-9463 office

(913) 424-8489 cell

APPENDIX B

TEACHER DEMOGRAPHIC INFORMATION

TEACHER DEMOGRAPHIC INFORMATION

1. Please enter your Teacher ID.

2. Please indicate your current grade level.
   - ○ 3rd
   - ○ 4th
   - ○ 5th
   - ○ 6th
   - ○ 7th
   - ○ 8th

3. Please indicate how many years have you been teaching.
   - ○ 1-4
   - ○ 5-10
   - ○ 11-15
   - ○ 16-20
   - ○ 21 plus

4. Please indicate your highest level of education.
   - ○ Undergraduate
   - ○ Masters
   - ○ Masters Plus 15
   - ○ Doctorate

5. Please indicate your gender
   - ○ Male
   - ○ Female

APPENDIX C

ADDITIONAL TEACHER RESPONSES

For the purpose of answering these questions, please define ***literacy*** as *reading, writing, listening, and speaking.*

1 10

Very poor Excellent

1. How would you rate the quality of your university preparation for ***literacy*** instruction?
   1 2 3 4 5 6 7 8 9 10

2. How would you rate the quality of the professional development, specific to ***literacy***, that you have received in your school district?
   1 2 3 4 5 6 7 8 9 10

3. Approximately how many professional-development sessions focusing on ***literacy*** and conducted by your school district have you attended in the past 2 years?
   0
   1-2
   3-5
   5-8
   More than 8

4. Have you ever participated in a book study with colleagues with the goal of discussing ***literacy***-related topics?
   Yes
   No

   If you answered 'yes,' please provide the following information:

   Topic or title of book:

   Number of times your study group met:

   Format used to discuss the book:

APPENDIX D

THE TEACHER SENSE OF EFFICACY FOR LITERACY INSTRUCTION SCALE

Teacher Beliefs - TSELI

This questionnaire is designed to help us gain a better understanding of the kinds of things that create challenges for teachers. Your answers are confidential.

Directions: Please indicate your opinion about each of the questions below by marking any one of the nine responses in the columns on the right side, ranging from (1) "None at all" to (9) "A Great Deal" as each represents a degree on the continuum.

Please respond to each of the questions by considering the combination of your current ability, resources, and opportunity to do each of the following in your present position.

		None at all		Very Little		Some Degree		Quite A Bit		A Great Deal
1.	To what extent can you use a student's oral reading mistakes as an opportunity to teach effective reading strategies?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
2.	To what extent can you use a variety of informal and formal reading assessment strategies?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
3.	To what extent can you adjust reading strategies based on ongoing informal assessments of your students?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
4.	To what extent can you provide specific, targeted feedback to students' during oral reading?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
5.	How much can you do to meet the needs of struggling readers?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
6.	To what extent can you adjust writing strategies based on ongoing informal assessments of your students?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
7.	To what extent can you provide your students with opportunities to apply their prior knowledge to reading tasks?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
8.	To what extent can you help your students monitor their own use of reading strategies?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
9.	To what extent can you get students to read fluently during oral reading?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
10.	To what extent can you model effective reading strategies?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
11.	To what extent can you implement effective reading strategies in your classroom?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
12.	To what extent can you help your students figure out unknown words when they are reading?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
13.	To what extent can you get children to talk with each other in class about books they are reading?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
14.	To what extent can you recommend a variety of quality children's literature to your students?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
15.	To what extent can you model effective writing strategies?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
16.	To what extent can you integrate the components of language arts?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
17.	To what extent can you use flexible grouping to meet individual student needs for reading instruction?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
18.	To what extent can you implement word study strategies to teach spelling?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
19.	To what extent can you provide children with writing opportunities in response to reading?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
20.	To what extent can you use students' writing to teach grammar and spelling strategies?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
21.	How much can you motivate students who show low interest in reading?	①	②	③	④	⑤	⑥	⑦	⑧	⑨
22.	How much can you do to adjust your reading materials to the proper level for individual students?	①	②	③	④	⑤	⑥	⑦	⑧	⑨

APPENDIX E

THE UNIVERSITY OF KANSAS HUMAN SUBJECTS APPROVAL

KU RESEARCH & GRADUATE STUDIES
The University of Kansas

7/12/2010
HSCL #18766

Jennifer Poggio
19005 West 98th Terrace
Lenexa, KS 66220

The Human Subjects Committee, Lawrence Campus (HSCL) has received your response to its expedited review of your research project

18766 Poggio/Barry (C & T) Relating Individual and Collective Teacher Efficacy Beliefs with Third, Fourth, Fifth and Sixth Grade Students' Reading Performance on the Kansas State Assessment in Reading

and approved this project under the expedited procedure provided in 45 CFR 46.110 (f) (7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

Since your research presents no risk to participants and involves no procedures for which written consent is normally required outside of the research context HSCL may waive the requirement for a signed consent form (45 CFR 46.117 (c) (2). Your information statement meets HSCL requirements. The Office for Human Research Protections requires that your information statement must include the note of HSCL approval and expiration date, which has been entered on the form sent back to you with this approval.

1. At designated intervals until the project is completed, a Project Status Report must be returned to the HSCL office.
2. Any significant change in the experimental procedure as described should be reviewed by this Committee prior to altering the project.
3. Notify HSCL about any new investigators not named in original application. Note that new investigators must take the online tutorial at http://www.rcr.ku.edu/hscl/hsp_tutorial/000.shtml.
4. Any injury to a subject because of the research procedure must be reported to the Committee immediately.
5. When signed consent documents are required, the primary investigator must retain the signed consent documents for at least three years past completion of the research activity. If you use a signed consent form, provide a copy of the consent form to subjects at the time of consent.
6. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.

Please inform HSCL when this project is terminated. You must also provide HSCL with an annual status report to maintain HSCL approval. Unless renewed, approval lapses one year after approval date. If your project receives funding which requests an annual update approval, you must request this from HSCL one month prior to the annual update. Thanks for your cooperation. If you have any questions, please contact me.

Sincerely,

Mary Denning

Mary Denning
Coordinator
Human Subjects Committee Lawrence

cc: Arlene Barry

Human Subjects Committee Lawrence
Youngberg Hall | 2385 Irving Hill Road | Lawrence, KS 66045 | (785) 864-7429 | Fax (785) 864-5049 | www.rcr.ku.edu/hscl

SS