DEVELOPMENT AND VALIDATION OF AN INSTRUMENT TO MEASURE SECONDARY TEACHERS' SELF-EFFICACY IN READING INSTRUCTION (STERI) ACROSS THE CONTENT AREAS

by

Erin Marie Ramirez

A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Education

Committee:

____________________ Chair

____________________ Chair

____

____

____________________ Program Director

____________________ Dean, College of Education and Human Development

Date: April 18, 2016

Spring Semester 2016
George Mason University
Fairfax, VA

Development and Validation of an Instrument to Measure Secondary Teachers' Self-Efficacy in Reading Instruction (STERI) across the Content Areas

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Erin Marie Ramirez
Master of Education
Rutgers University, 2011
Bachelor of Arts
Syracuse University, 2005

Director: Angela Miller, Assistant Professor; Michelle Buehl, Associate Professor
College of Education and Human Development

Spring Semester 2016
George Mason University
Fairfax, VA

Dedication

“Only those who will risk going too far can possibly find out how far one can go.”
-T.S. Elliot

To my immediate family…

My mother, for your unwavering support as I trudged through yet another degree. You truly have been my Mama Bear and have never stop pushing me to be the best version of myself. You have helped make this possible and I will never forget everything you have done for me. “All that I am or ever hope to be, I owe to my mother.”

Dad, for always believing in me and pushing me to be better. Viva Colombia!

The men in my life: Andrew, Alex, and Adam, for always supporting me even when you had no idea what my research or job entailed. For the hours of Mario Kart, Disney Monopoly, Settler of Catan, and Exploding Kittens; you helped me forget about the mountains of work awaiting me when I got home from vacation. Let’s go Jets!

Nanny Irene, you have inspired me my entire life. I hope this makes you and Poppy proud; even though I still can’t write you prescriptions. I can only hope to live as long and full a life as you; ninety has never been so funny.

Lucky, the world’s greatest dog, for being my favorite little brother; miss you always.

To my extended family…

My cousin Marissa who has called me “docta” since the day I got accepted. Remember that with hard work, anything is possible.

The Rye’s for all the Walt Disney World vacations that helped solidify that family always sticks together, supports one another, and celebrates milestones together.

The Rodriguez, Ramirez, Labinsky, and Hieronymus gangs for your steadfast support as I pursued my lifelong goal of getting a PhD. Everyone deserves a family that is as wonderful, caring, supportive, and loving as you are. I appreciate each and every one of you more than words can say.

To the family I have built through friendships…

Lesley King, my work wife, for all of your encouragement as I pursued my PhD and worked through my dissertation. You are the most caring, loyal, unselfish, loving, thoughtful, and wonderful person I have ever met. You helped me get through this program and I will forever be grateful for everything you have done for me. You will always be my number one work wife, and I cannot wait to celebrate you as Dr. King.

Liz and Dan, my BFFs, for helping distract me from my work with Halloween parties, Sundays with football and Chinese food, and weekend trips to Philadelphia. You will never know how much these events helped me stay sane throughout this process.

The Boonies Crew (Liz, Nick, & Keri) for always reminding me of where I came from and helping me stay humble. We have come a long way since seventh grade.

Leila Richey Nuland, for being my first mentor when I got to Mason and helping me pursue my own research. You will forever be my conference roommate.

Paula and John, for being my second set of parents and always believing in me. Your support has meant the world to me.

Anna, Lucas, Sonia, and Zachary for being the most amazing little humans. You are the future, never forget to follow your dreams, and I can't wait to see all you accomplish in your lives.

To my best friend…

Sarah, for being with me every step of the way from applying to graduating. You were there through all the tears, screams, milestones, and accomplishments; without you, none of this would have been possible. Whenever I started to waver, or feel like I made a mistake, you would reel me back in and remind me why I was pursuing these goals; thank you for being my anchor and partner throughout this and everything else life has thrown our way. Always and Forever, in cursive, Mary Poppins.

I am nothing without my family and friends, this is for each one of you who has shown me that kindness, compassion, and support can help people move mountains.

Acknowledgements

"There is a vitality, a life force, an energy, a quickening that is translated through you into action, and because there is only one of you in all of time, this expression is unique. And if you block it, it will never exist through any other medium and it will be lost. You have to keep yourself open and aware to the urges that motivate you. No artist is pleased. [There is] no satisfaction whatever at any time. There is only a queer divine dissatisfaction, a blessed unrest that keeps us marching and makes us more alive than the others." -Martha Graham

First and foremost, to my chairs, Dr. Angela Miller and Dr. Michelle Buehl, thank you for all of the time, energy, and support you gave me throughout my time at Mason. Through all the meetings, drafts, questions, missteps, amendments, and occasional tears you never stopped supporting, advising, and encouraging me. You have been the best advisors and mentors, and through your guidance I have become a better writer, educator, scholar, mentor, and person. At times you pushed me beyond my limits and for this I am grateful, for we never know what we can really accomplish until we have been pushed farther than we thought possible.

To my committee, Dr. Gary Galluzzo and Dr. Elizabeth Sturtevant, thank you for your support, commitment, and advice throughout this process. You always sought to make me feel comfortable at Mason and gave me the opportunities to thrive. I will never forget your dedication to teaching, mentoring, and advising even after becoming prolific in the educational field. I can only hope to accomplish all you two have during my time in academia. Thank you, and Go Orange!

To my mentors at Mason and beyond, Dr. Seth Parsons, Dr. Allison Ward Parsons, and Dr. Ana Taboada Barber, thank you for providing me with amazing research and writing opportunities that helped contribute to my success as a scholar. Thank you for putting me on your shoulders and helping me succeed. I will forever be grateful for the support, encouragement, and guidance I received from each one of you.

To Aysha Puhl, Jenn Johnson, and Jackie Nash, thank you for giving me my first opportunity to teach in higher education. You opened the door and I have never stopped looking back. Project Peak forever!

To the experts who helped review my measure, Dr. Anita Woolkfolk Hoy, Dr. Ellen Usher, Dr. Nancy Frey, and Dr. Maryann Mraz, thank you for your support, kindness, time, and advice. I look up to each and every one of you, and I can never thank you enough for all you did to help me through this process.

To my fellow students who helped review a very early, and long, version of my measure, thank you for supporting me and providing feedback to help make this project a success.

To all the teachers across the United States who took my measure and helped me to get the sample size I needed. I appreciate all of your help in this endeavor.

To Joan Stahle, thank you for everything you do for the PhD program and Mason. You were always a source of information and knew the answer to any question I could ever ask, and I will forever appreciate all of your guidance throughout my time at Mason.

To all my former teachers, especially Mrs. Lee Ann Smith Kubbishun and Mrs. Nancy Terhune, thank you for being my first models of true teaching. You have been my teaching mentors since seventh grade, and I still carry with me the fondest memories of being in your classes. Thank you for giving me models to strive for and the support to chase my dreams.

To all my students at Mason, thank you for helping me become a better educator. You revitalized my drive and dedication to the craft of teaching, and for that I will always be grateful.

To the Starbucks employees at the Burke and Fair Lakes locations, thank you for not judging me. You never stared at me funny as I spent twelve hours at the corner table writing, ordering drinks every few hours, or pacing when I couldn't think of what to write next. And when I did the same thing the next day, you smiled and asked if I still wanted to add that extra shot; thank you.

Finally, to the creators, writers, and actors of The Office, thank you for making me laugh even when I felt like I would never make it through this program. Getting a PhD can be one of the most isolating experiences, but watching your show helped me to laugh through tears and smile through frustration. Thank you for being my comic relief in a world full of chaos.

I can sum up this entire experience with one quote, "Sometimes I'll start a sentence and I don't even know where it's going. I just hope to find it along the way." –Michael Scott

Table of Contents

List of Tables

List of Figures

Definition of Key Terms

Adolescent Literacy: The literacy development and instructional needs for students in grades 4-12 (Biancarosa & Snow, 2006).

Content Area Literacy: The level of reading and writing skill that learners need in an academic subject to comprehend and respond to ideas in texts used for instructional purposes (Vacca, 2002).

Disciplinary Literacy: Advanced literacy instruction embedded within content area classes that emphasizes literacy strategies specific to each content area (Shanahan & Shanahan, 2008).

Professional Development: Learning opportunities available to teachers and other educational personnel with the goal of strengthening understanding and skills associated with their teaching practice (Darling-Hammond et al., 2009).

Reading Motivation: Individual's goals and beliefs with regard to reading. Reading motivation then may influence the individual's activities, interactions, and learning with text (Guthrie & Wigfield, 1999).

Reading Strategies: Deliberate, goal-directed attempts by students to control and modify their efforts to decode text, understand words, and construct meanings of text (Afflerbach, Pearson, & Paris, 2008).

Self-Efficacy: Beliefs in one’s capabilities to organize and execute the courses of action required to produce given attainments within specific contexts (Bandura, 1977).

Struggling Readers: Students who are two or more grade levels behind in reading. These struggling readers lack the ability and/or fluency to facilitate comprehension; others can more accurately read, yet they lack the strategies to aid in comprehension of texts; and still others have the strategies to help with comprehension, but they do not know how to use these across content-area classrooms (Biancarosa & Snow, 2006).

Abstract

DEVELOPMENT AND VALIDATION OF AN INSTRUMENT TO MEASURE SECONDARY TEACHERS' SELF-EFFICACY IN READING INSTRUCTION (STERI) ACROSS THE CONTENT AREAS

Erin Marie Ramirez, Ph.D

George Mason University, 2016

Dissertation Director: Dr. Angela Miller and Dr. Michelle Buehl

The purpose of this study was to develop and validate a scale to measure Secondary Teachers Sense of Efficacy for Reading Instruction (STERI) across the content areas. The study began by examining the adolescent literacy and teacher self-efficacy literature. Following social cognitive theory and Bandura's (1977) conception of self-efficacy, items were written around three hypothesized factors that encompassed adolescent literacy instruction. The three hypothesized factors were: Self-Efficacy for Using Content-Area Reading Strategy Instruction, Self-Efficacy for Motivating Students to Read, and Self-Efficacy for Teaching Struggling Readers.

Following Gehlbach and Brinkworth's (2011) framework, face and content validity was established using two rounds of expert review (*n* = 10) and cognitive interviews (*n* = 12). After all initial revisions were completed, Exploratory Factor Analyses (EFA) was conducted on the 31-item STERI instrument. EFA revealed a three

factor structure that included 22 of the original 31 items and accounted for 74.95% of the variance equally across the three factors (Factor 1, 28.26%; Factor 2, 25.69%; Factor 3, 24.32%). The items factored onto the three hypothesized factors of self-efficacy for: Using Content-Area Reading Strategy Instruction (n = 7), Motivating Students to Read (n = 7), and Teaching Struggling Readers (n = 8). Additionally, reliability was established for each factor and the overall scale using Cronbach's alpha (Factor 1, α = .96; Factor 2, α = .95; Factor 3, α = .95; Overall, α = .98). Finally, convergent and discriminant validity was established through correlational and Analysis of Variance analyses.

Based on the findings from all analyses, the STERI was shown to be a reliable and valid instrument in its initial examination. Thus, further research should be conducted to confirm the factor structure, reliability, and validity of the STERI. Upon confirmation of the factor structure of the STERI, the instrument can be used as a diagnostic tool for literacy professional development in an attempt to increase teachers' sense of efficacy for reading instruction at the secondary level. By measuring teacher self-efficacy specifically an attempt can be made to causally link teacher self-efficacy, teacher instruction, and student achievement.

Chapter One: Introduction

Adolescent literacy has become a prevalent concern in the United States as 64% of 8th grade students are reading at basic and below basic levels, and 70% of high school students are considered struggling readers who read at least two grade levels behind, and are in need of extensive remediation (Biancarosa & Snow, 2006; Nation's Report Card, 2013). As literacy encompasses the actions of reading, writing, speaking and listening, in order to address the issues of low reading ability in the US, the current study addressed adolescent literacy instruction from the perspective of reading. As secondary schools prepare students to become college and career ready, it becomes paramount that adolescent students can read at the appropriate level to be successful in these college and career settings.

Research has shown that students of all ability levels can improve their comprehension when reading instruction is given in the secondary content areas (Fisher, Frey, & Nelson, 2012; Nokes & Dole, 2004; Underwood & Pearson, 2004). However, many secondary teachers have little to no training on how to incorporate literacy instruction in their content area; leading to differences in how/if literacy instruction is infused into the secondary classroom (Fisher & Ivey, 2005). This lack of training leads many teachers to have little confidence (i.e., low self-efficacy) in their ability to successfully integrate reading instruction in their classroom and in turn adverse beliefs towards secondary literacy instruction (Hall, 2005; O'Brien, Stewart, & Moje, 1995).

These self-efficacy beliefs influence instruction (Guo, Piasta, Justice, & Kaderavek, 2010; Tschannen-Moran, Hoy & Hoy, 1998) which impacts student achievement (Ross, 1992); thus it becomes important to measure teachers' sense of efficacy in relation to reading instruction. Ascertaining teachers' sense of efficacy can offer insights into how successful teachers feel they are at incorporating reading instruction in their content area. Diagnosing a teacher's sense of efficacy in literacy instruction can lead to the development of professional practices that increase content-area reading instruction and consequently improve adolescent reading achievement. Currently, no self-efficacy measures exist that examine secondary in-service teachers' sense of efficacy in reading instruction across the content areas in a context (e.g., secondary), domain (i.e., content or subject area), task (e.g., reading instruction) specific way with an emphasis on this target population (i.e., in-service). Thus, the purpose of the current study is the development and validation of a teacher self-efficacy instrument to measure secondary (context) content area (domain) teachers' sense of efficacy in reading instruction (task) using in-service teachers (target population).

Albert Bandura (1977) defined self-efficacy as one's beliefs about one's ability to execute a specific task and achieve a specific outcome. Furthermore, Bandura (1986) articulated that self-efficacy is a context, domain, and task specific construct, and thus should be measured in this manner. In relation to teaching, teachers' self-efficacy beliefs may change based upon the context (i.e., elementary or secondary), domain (i.e., content or subject area) and the task (i.e., specific teaching task) in which they are teaching. Bandura defined self-efficacy as context, domain, and task specific, yet the most commonly used teacher self-efficacy instruments (i.e., Teacher Efficacy Scale [TES], Gibson & Dembo, 1984; Teacher Sense of Efficacy Scale [TSES], Tschannen-Moran & Woolfolk Hoy, 2001) assess teachers' self-efficacy in a general manner. In order to correlate teacher self-efficacy to specific outcomes, both variables must be measured specifically in order to examine explicit connections (i.e., if student achievement is measured task specifically [reading achievement], than teacher self-efficacy must also be measured task specifically [self-efficacy for teaching reading]).

With the development of new context, domain, and task-specific measures, researchers can more explicitly link teacher self-efficacy with student achievement, and further examine other factors (e.g., track of students) that may impact teachers' self-efficacy. Additionally, self-efficacy measures can be used to help determine what makes effective and ineffective teachers (Bandura, 1997; Gist & Mitchell, 1992; Wheatley, 2005). Thus, if a literacy-specific measure was created, researchers could determine factors and variables that make secondary teachers more or less effective in utilizing literacy instruction in their content area classrooms. Moreover, researchers could examine what factors influence teacher's low self-efficacy to attempt to heighten these beliefs through professional development and other professional opportunities.

Adolescent Literacy Instruction

With the rise of technology and digital texts, adolescent readers are faced with the challenges of engaging, interpreting, and synthesizing both print and non-print texts of increasing complexity to make meanings of their educational and social worlds (International Reading Association, 2012; Luke & Elkins, 2000; Vacca, 1998). As text complexity increases with the progression through the secondary school system, more and more students continue to struggle with the demanding nature of adolescent literacy. These struggles were further noted in The Nation's Report Card in Reading (2013) which delineated, "22% of eight graders scored at a Below Basic level, 42% were at a Basic level, 32% scored at a Proficient level, and 4% scored at an Advanced level" (National Center for Educational Statistics, 2013, p. 7) and has also been demonstrated consistently through the literature (Biancarosa & Snow, 2006; Heller & Greenleaf, 2007; Snow & Moje, 2010). Thus, teachers at the secondary level are consistently faced with students who are not achieving at grade level in literacy; leading to the necessity for secondary content area teachers to provide reading instruction in all content areas to prepare these students to be college and career ready.

Content-Area and Disciplinary Literacy

Consequently, with the creation and adoption of the Common Core State Standards (CCSS), policymakers have declared that cross-disciplinary literacy is now essential in all levels of K-12 education (National Governors Association, 2010). Under the CCSS, all teachers must now teach literacy in their classroom, regardless of their educational background and experiences with literacy coursework in their preparation programs. As literacy instruction is now a necessary component of secondary education, teachers must feel they can effectively use literacy instruction within their secondary classrooms in order to be successful. If teachers do not feel they can effectively use literacy instruction in their classroom then it is likely they will not provide this instruction, and adolescents will continue to perform at below proficient levels.

As a result of the increasing number of adolescent students labeled as struggling readers, reading at least two grade levels behind (Biancarosa & Snow, 2006), researchers have developed methods by which content area teachers can infuse literacy instruction into their classrooms. Within the literacy community, two approaches (i.e., content-area literacy and disciplinary literacy) have emerged that focus on the direct instruction of reading strategies as an anchor for secondary content area literacy instruction. Both approaches utilize explicit strategy instruction as a way for teachers to motivate student to engage with literacy tasks to improve their comprehension and proficiency as readers. As proficiency increases, students become more engaged in literacy, motivated to read, and improved student achievement occurs (Irvin, Meltzer, & Dukes, 2007). Where the two approaches differ is on the type of reading strategy instruction teachers are engaged in with students.

Research on content-area literacy demonstrates how a consistent set of reading strategies (e.g., using background knowledge) can be used by students across content area classrooms to improve reading comprehension; especially for struggling readers as it offers them a strategic approach to reading across the content areas (Gersten, Fuchs, Williams, & Baker, 2001; Vacca, 2002). With this approach, secondary teachers instruct students on the definition and application of these generic strategies, then aid students in implementing these strategies within their classroom, with the end goal of improving student comprehension. Further, if students are able to successfully use content-area literacy strategies, their comprehension improves leading to an increase in reading engagement and subsequent motivation to read (Guthrie & Wigfield, 2000). On the other hand, disciplinary literacy is the practice of teaching students specific disciplinary reading strategies (e.g., thinking like a scientist) to use within a specific discipline (Shanahan & Shanahan, 2008). Additionally, disciplinary literacy has been defined as an advanced form of literacy (Shanahan & Shanahan, 2008). Within disciplinary literacy instruction, each secondary teacher would be responsible for teaching students how to utilize specific strategies within their discipline, in order to improve reading comprehension.

When discussing content-area and disciplinary literacy, the fundamental difference in approaches creates the argument to first use content-area then disciplinary literacy. One critical difference is the type of strategies used to increase student comprehension (i.e., generic versus discipline-specific strategies). Another difference is that content-area literacy is an intermediate form of literacy, whereas disciplinary literacy is an advanced form of literacy (Shanahan & Shanahan, 2008). Seeing these two approaches as falling onto a literacy continuum, content-area literacy occurs first (i.e., intermediate form of literacy) and then students move into disciplinary literacy (i.e., advanced form of literacy). Additionally, empirical research has demonstrated that content-area reading strategies improve student's reading achievement (i.e. Fisher, Frey, & Williams, 2002); however this evidence has not yet been demonstrated for disciplinary literacy instruction. Thus, this study focuses on the creation of a literacy-specific teacher self-efficacy measure centered on the tenets of content-area literacy.

Within the content-area literacy framework, three central ideas were explored: reading strategy instruction (Fisher & Frey, 2011; Fisher et al., 2012), motivating students to read (Guthrie & Wigfield, 1999), and struggling readers (Marzano, Pickering, & Pollock, 2001. These three factors were determined to be at the center of the content-area literacy literature as fundamental attributes to successful adolescent reading instruction. Given that research has demonstrated that teacher beliefs' influence instructional practices (Guo et al., 2010; Tschannen-Moran et al., 1998), and instruction practices impact student achievement (Ross, 1992), it is important to measure teachers' beliefs in relation to literacy instruction to see how these beliefs detract or enhance instruction and subsequent students' achievement. One specific belief, that research has shown to be correlated with teacher instruction, is teachers' beliefs about their effectiveness (self-efficacy) with a given context, domain, and task.

Teacher Self-Efficacy

The field of self-efficacy has been guided by two theories, Rotter's (1966) locus of control and Bandura's (1977) social cognitive theory, which have predominately influenced the construct of self-efficacy and its subsequent study. As self-efficacy is a latent construct (i.e., not directly observable), a theoretical framework must be chosen and consistently adhered to when studying the construct and devising instrumentation. Using Rotter's locus of control theory, teachers' self-efficacy would be influenced by internal versus external factors; thus, factors the teachers could control (internal) versus things outside of the teachers control (external). Bandura's social cognitive theory focuses on triadic reciprocation where the personal, behavioral, and environmental factors all influence each other in a reciprocal fashion. Additionally, Bandura (1986) defined self-efficacy as being context, domain, and task-specific. Bandura's social cognitive theory (i.e., triadic reciprocation between person, behavior, and environment) and definition of self-efficacy (i.e., context, domain, and task-specific) describe the construct in a manner that is congruent with my conception of the self-efficacy, instead of the two-pronged approach (i.e., internal versus external) stated in locus of control theory.

Thus, for the purposes of this review, only Bandura's (1977) definition of self-efficacy was used as the current study aligns with self-efficacy from a social cognitive perspective (i.e., multi-faceted; context, domain, and task-specific) instead of a locus of control perspective (i.e., external vs. internal). Furthermore, the TES and other measures structured around Rotter's locus of control framework often incorrectly measure self-efficacy and outcome expectancy as highly correlated; however, this contradicts Bandura's (1986) conception that the two constructs are distinguishably different. Therefore, only teacher self-efficacy research that utilized scales developed based on Bandura's framework was presented when providing rationale for the study of teacher self-efficacy as social cognitive theory is the framework from which the current study and instrument are aligned.

Self-efficacy is the beliefs a person holds about one's ability to perform a certain task that influence how well that person will then perform that task (Bandura, 1977; Pajares, 1996). Hence, how effective a person believes one can be at a specific task will in turn impact how well one will perform at that task. Further, Bandura (1986) described how self-efficacy beliefs are the driving force behind one's motivation, confidence level, and ability to perform well. Thus, even if a situation is extremely difficult, self-efficacy beliefs will influence motivation, confidence, and ability above and beyond a person's actual ability level. That is, how a person feels about personal abilities is more impactful than abilities themselves. In teaching, even if a teacher possesses the content and pedagogical knowledge to be successful, the belief he/she has about his/her effectiveness (i.e., self-efficacy) can either enhance or inhibit his/her teaching ability.

After his original conception of self-efficacy in 1977, Bandura explored self-efficacy beliefs further and described the four sources of information that impact one's sense of efficacy: mastery experiences, vicarious experiences, social persuasion, and physiological state. For teachers, the teaching experiences a teacher has (mastery experiences), the modeling/instruction a teacher receives (vicarious experiences), the feedback/guidance a teacher hears (verbal persuasion), and the level of emotion a teacher feels (physiological state) all interact and contribute to a teacher's overall sense of teaching efficacy. Additionally, it has been hypothesized that mastery experiences influence one's sense of efficacy above the other three sources (Bandura, 1997). Thus, in teaching, the experiences a teacher has in the classroom may impact his/her confidence, motivation, and instruction more than the other three sources of self-efficacy beliefs (Tschannen-Moran et al., 1998).

Specificity of Self-Efficacy Beliefs

According to Bandura's (1986) social cognitive framework of self-efficacy, a person's perceived self-efficacy is context, domain, and task-specific. For teachers, their context (i.e., elementary or secondary), domain (i.e., content or subject area), and the specific task to which they are engaged (e.g., content-area literacy instruction) will change how efficacious they feel. In order to more accurately assess teachers' sense of efficacy, new measures should be created that assess teachers' sense of efficacy in context, domain, and task specific ways in order to more accurately represent the construct of teaching self-efficacy. Moreover, Bandura's (1986) social cognitive framework of self-efficacy hypothesized that mastery experiences may influence one's perceived self-efficacy more than any of the other forms of persuasion; therefore, measures should be designed to assess the specific experiences of the target population (i.e., pre-service or in-service teachers). Thus, new measurement must address teacher self-efficacy from a context, domain, and task-specific way, with attention paid to the mastery experiences of the target population.

Current study of self-efficacy. Currently, little is known about secondary content area teachers' self-efficacy beliefs about literacy instruction in the content areas (Cantrell & Callaway, 2008; Cantrell & Hughes, 2008) as none of the current teacher self-efficacy instruments are focused specifically on content-area literacy instruction at the secondary level. In their review of teacher self-efficacy research from 1998-2009, Klassen, Tze, Betts, and Gordon (2011) found that 60% (n = 130) of studies measured general teaching efficacy, with 15% (n = 32) of the studies focused on science; whereas only 2% (n = 4) of the 218 studies they examined centered on the domains of teaching language and literacy. As the need for literacy instruction at the secondary level continues, it becomes necessary to measure secondary teachers' sense of efficacy in literacy instruction, to more accurately link these practices to student outcomes in an effort to decrease the number of struggling readers.

Bandura (1997) argued the increase in studies focused on science teachers' self-efficacy was due to the shifting nature of society; however, the same causal argument can now be made for an increased concern of literacy instruction in the secondary classroom. With the emergence of the CCSS, the increased number of students who are classified as struggling readers, and those that read at basic and below basic levels, the transformative nature of education has shifted towards an increase of literacy instruction across all classrooms. These societal changes have forced adolescent literacy instruction into content area classrooms, making researchers find ways to measure teachers' sense of effectiveness in infusing literacy instruction into their content area to improve student's reading achievement.

Bandura (1997) argued that concern over teacher self-efficacy fluctuates with societal interests in student achievement. Therefore, as adolescent reading achievement continues to be an issue in the US (e.g., 64% of 8^{th} grade students reading at basic or below basic levels; Nations Report Card, 2013), the creation of literacy self-efficacy measures has subsequently increased. These measures have been designed to focus on various aspects of literacy instruction: pre-service teachers' sense of efficacy in reading instruction (i.e., Reading Teachers' Sense of Efficacy Scale [RTSES], Haverback, 2007; Haverback & Perault, 2011), teachers' sense of efficacy in literacy instruction (i.e., Teachers' Sense of Efficacy in Literacy Instruction [TSELI], Tschannen-Moran & Johnson, 2011), and teachers' sense of efficacy in the teaching of the English Language (i.e., The English Language Teaching Efficacy Instrument [ELTEI], Akbari & Tavassoli, 2014; Exceptional Children who are English Learners [EXCEL] Teacher Inventory, Paneque & Barbetta, 2006). Although an increase in literacy-specific measures is an advancement, none of these measures address the context of secondary schools, the task of content-area literacy instruction, and the target population of secondary in-service teachers.

Summary

Perceived self-efficacy is context, domain, and task-specific, and as it is hypothesized, mostly highly influenced by one's mastery experiences (Bandura, 1977, 1986). Therefore, teacher self-efficacy measures should encompass the specific context (i.e., elementary or secondary) domain (i.e., content or subject area), task (e.g., content area reading instruction), and mastery experiences (i.e., pre-service or in-service teachers), with which the researcher is interested in studying. In order to assess secondary teachers' sense of efficacy in reading instruction, a new measure must be created that encompasses in-service (mastery experiences of target population) secondary teachers' (context) sense of efficacy in reading instruction (task) across the content areas (domain).

Furthermore, to make the newly created measure as task-specific as possible, it becomes necessary to define specific aspects of content-area reading instruction with which to focus. Thus, according to the adolescent literacy research, three specific areas (factors) were defined: reading strategy instruction, motivating students to read, and struggling readers. Therefore, assessing teachers' sense of efficacy for infusing these three tasks into their content area classrooms becomes a primary concern. The goal of the current study was to: articulate the need for a self-efficacy instrument to measure secondary teachers sense of efficacy in reading instruction (STERI) across the content areas, describe the creation of the instrument, discuss validity and reliability, and explore the factor structure of the measure.

Chapter Two: Literature Review

The purpose of this chapter was to: demonstrate the need for content-area literacy instruction at the secondary level, depict issues and gaps within current teacher self-efficacy instrumentation, and delineate the need for new teacher self-efficacy measurement focused on secondary teachers' sense of efficacy in reading instruction (STERI) across the content areas. This commenced by first discussing the history of adolescent literacy, the need for literacy instruction in the secondary classroom, and the argument for content-area literacy instruction across all content areas. Furthermore, this chapter describes the distinction between generic reading strategy use across content areas (content area literacy) and specific reading strategy use for each content area (disciplinary literacy); providing evidence to support the inclusion of content-area literacy practices for the purposes of this study.

Next, this chapter explored the history of teacher self-efficacy research and how two distinct theoretical frameworks have influenced the construct: Rotter's (1966) locus of control and Bandura's (1977, 1986) social cognitive theory. Then, this chapter examined how self-efficacy impacts instruction and educational outcomes based upon current research in the field. Additionally, issues and gaps with the current generic and domain-specific instruments were discussed. The need for new measurement in the field of teacher self-efficacy and content-area literacy instruction is addressed. Finally, the hypothesized factors included in the creation of the STERI were delineated and described in detail.

Search of Existing Literature

A literature review was conducted using the Social Sciences Citation Index, APA Psycnet, Education Full Text, and ERIC databases using the search terms: social cognitive theory, teacher self-efficacy, content specific teacher efficacy measures, content area literacy, disciplinary literacy, adolescent literacy, measure development, secondary teachers' self-efficacy, teachers' self-efficacy and literacy instruction, literacy professional development, and literacy teachers' self-efficacy beliefs. Additionally, searches were conducted in relevant educational journals (e.g., *Teaching and Teacher Education, Journal of Educational Psychology*).

Specific parameters were used in selecting articles for inclusion in this chapter. For instance, only teacher self-efficacy studies were included, collective efficacy was not considered within the scope of this study. Only peer-reviewed articles were used in the final literature review. For the purposes of this review, just studies that utilized teacher self-efficacy measures based off of Bandura's (1977,1986) social cognitive framework were included when discussing teacher self-efficacy, instruction, and outcomes. Additionally, only measures that were developed based on Bandura's social cognitive theory (e.g., Teacher Sense of Efficacy Scale [TSES], Tschannen-Moran & Woolfolk Hoy, 2001) were discussed in detail. All studies in which Rotter's (1966) locus of control theory was used to measure teacher self-efficacy were not included in the discussion of teacher self-efficacy and outcomes. Finally, measures in which Rotter's framework of self-efficacy was used for development (e.g., Teacher Efficacy Scale [TES], Gibson & Dembo, 1984) are not discussed in detail.

Adolescent Literacy

Adolescent readers have a much different task than their younger counterparts, and therefore adolescent literacy must be studied separately from early literacy. Adolescent literacy encompasses more complex and demanding skills in which adolescent readers must make meaning of both print and non-print texts, while trying to gain knowledge from what they are reading (Vacca, 1998). Additionally, adolescent readers must be able to read, comprehend, interpret, and critique complex discipline-specific literature across different contexts (International Reading Association, 2012; Moje, Overby, Tysvaer, & Morris, 2008). As technology has advanced and the realm of adolescent literacy has changed, more and more attention has been paid to studying adolescent literacy instruction across the content areas.

Although many definitions of adolescent literacy emerged over the years, the focus never shifted from the complex nature of adolescent literacy and the multitude of skills required for adolescents to be effectively literate. The difficult nature of adolescent literacy has become even more apparent in today's educational system with the adoption of the Common Core State Standards (CCSS) in 2010. The CCSS state that reading, writing, speaking, and listening are integral parts of every subject from K-12 (National Governors Association, 2010). Therefore, all secondary students are expected to be literate in all subject areas.

History of Reading Instruction in Secondary Education

Reading instruction in the content areas was promoted by Huey in 1908 and many attributed his book, "The Psychology and Pedagogy of Reading" as the first and seminal piece on "how we read and learn to read" across all content areas (Kamil & Bernhardt, 2009). As the conception of reading instruction across the content areas advanced, studies encouraged the use of reading-to-learn instruction regardless of content areas (Thorndike, 1917). Within this framework, Thorndike (1917) described how reading was thinking, and involved more than just reading words on a page. He went on to describe that reading involved synthesis, analysis, problem solving, inferencing, generalizing, and concentrating as a means to learn specific content (Thorndike, 1917).

Following in the footsteps of Huey and Thorndike, Gray (1925) took up the cause of adolescent reading instruction across the content areas and argued that every teacher was a teacher of reading (Moore, Readence, & Rickelman, 1983); a concept which has since be re-envisioned by many researchers (e.g., Fisher & Ivey, 2005; Shanahan & Shanahan, 2008). In 1978, Herber's publication of *Teaching Reading in the Content Areas*, gave the field a structured guide that discussed teaching reading processes and content matter simultaneously in all content areas. This expanded the field of adolescent literacy, and researchers have since supported the idea that adolescent students benefit from literacy instruction in all content area classes (Alvermann, Phelps, & Ridgeway, 2007; Biancarosa & Snow, 2006; Conley & Hinchman, 2004; Hall, 2005; Kamil, 2003; Moore et al., 1983).

Within the literature on adolescent literacy instruction, two different approaches have formed (e.g., content-area literacy and disciplinary literacy) that focus on reading strategy instruction to increase adolescent literacy achievement. According to some scholars (Fisher & Frey, 2008 for example) content-area literacy focuses on the use of a generic set of reading strategies (e.g., activating background knowledge) that students learn and use across the content areas. Whereas disciplinary literacy centers on the use of discipline-specific strategies (e.g., learning to read like a historian) that students learn and use in each specific discipline (Shanahan & Shanahan, 2008 for example). As the literacy community continues to disagree on what "best practices" are for literacy instruction, American adolescents have shown no steady increase in reading scores over the past twenty years (Grigg, Donahue, & Dion, 2007).

In 2013, The Nation's Report Card in Reading delineated that 64% of students were reading at basic or below basic levels, and Biancarosa and Snow (2006) reported that 70% of high school students are identified as struggling readers who are two or more grade levels behind in reading and in need of extensive remediation. Although high stakes tests and grade level equivalents are not the only measures of reading comprehension, the research community has agreed that many adolescent students are not reading at grade level, and therefore are not being prepared to be college and career ready (e.g., Biancarosa & Snow, 2006; Fisher & Frey, 2005; Grigg et al., 2007). Thus, it becomes important for secondary teachers to have the instructional capability and effectiveness to infuse literacy instruction across the content areas.

One of reasons students struggle in reading is the difficult and often changing nature of texts as they proceed through school (Duke, 2000; Lee & Spratley, 2010). Students must have the ability to read, comprehend, interpret, analyze, and engage with text of varying purpose, text structure, style and vocabulary in both print and non-print forms (Alvermann, 2001; Bean, 2000; Heller & Greenleaf, 2007; Herber, 1978; Moje, 2000; Vacca, 1998). Although text complexity and variety increases, formal reading instruction ends in the sixth grade in many schools, with no one department responsible for literacy instruction/achievement (May, 2007). Most often, the responsibility for literacy instruction falls to the English department; yet many English teachers have little to no formal training in reading instruction (The National Council of Teachers of English, 2006). To improve the literacy skills of adolescent readers, literacy instruction at the secondary level must come from teachers across content areas.

Content area literacy. As adolescents work to comprehend information from a variety of texts, content-area literacy skills become essential to aid in students' successful interpretation and critical examination of texts across all content areas. According to Fisher and Ivey (2005), content-area literacy focuses on the use of literacy activities (e.g., reading, discussions) to help students construct and co-construct knowledge from a variety of perspectives. Additionally, content-area literacy focuses on the ability to use reading and writing in a particular content area (Vacca, 2002), and to read, interpret, and critique a variety of texts within a given domain (Bain, 2006). Adolescent literacy researchers agree that, regardless of subject area, all teachers should be able to assist students in reading to comprehend and make meaning of text in their domain or content area (Duke, 2000). Instruction on how to read and comprehend in the content areas is essential for students to be successful at the secondary level (Duke, 2000; Moss, 2005; Wood, 2003). Therefore, all teachers must instruct students on how to use reading strategies to improve literacy skills, increase comprehension, and further develop the building of knowledge among adolescents.

Furthermore, researchers have demonstrated that content-area literacy instruction is correlated with an increase in students' ability to comprehend and build knowledge from text across the curriculum (Biancarosa & Snow, 2006; Fisher et al., 2002; Kamil, 2003; Moore et al., 1983; Vacca, 2002). For example, Fisher and colleagues (2002) found that with sustained professional development structured around content-area literacy instruction, teachers were able to incorporate content area reading strategies into their classrooms, and students' reading achievement increased by two grade-level equivalents on a standardized reading assessment (i.e. Gates-MacGinitie).

Key areas in content area literacy. In order to address the number of struggling readers in the US who continue to read at least two levels below grade level, targeted interventions, and professional development opportunities have arisen that focus on literacy instruction in the secondary classroom. Researchers have addressed specific aspects of adolescent literacy instruction that have been shown to improve reading achievement for both struggling and more proficient readers (e.g., Fisher et al., 2012). Two such areas are cognitive reading strategies and motivating students to read.

Cognitive reading strategies. Content-area literacy proponents articulate that cognitive reading strategies (e.g., drawing connections) can traverse content areas

(Biancarosa & Snow, 2006), improve adolescent literacy achievement (Fisher et al., 2002), create a strategic approach to reading for struggling readers (Vacca, 2002), and motivate students to read (Guthrie, Wigfield, & VonSecker, 2000). Cognitive reading strategies are goal-directed actions to control and modify a reader's efforts to decode, understand, and make meaning of texts (Afflerbach, Pearson, & Paris, 2008). These reading strategies are used across content areas and many scholars believe that all teachers should engage in content-area reading strategy instruction (Ness, 2007).

Additionally, content-area reading strategy instruction has been shown to improve adolescent reading achievement (Fisher et al., 2002). Further, utilizing cognitive reading strategies in content area classrooms offers struggling readers explicit strategic actions and plans (e.g., summarizing key points, determining main idea and supporting details) to engage in while reading which helps build comprehension (Pardo, 2004). Finally, the use of cognitive reading strategies in content classrooms creates an environment where students build competence and become aware of that competence which further motivates them to engage in reading practices (Guthrie et al., 2000).

Using cognitive reading strategy instruction at the secondary level can create a supporting learning environment for struggling readers as this explicit strategy instruction has been linked to increased student comprehension (e.g., Anderson, Chan, & Henne, 1995). Although many teachers assume that students will naturally develop the ability to comprehend text (Denton & Fletcher, 2003), studies have shown that comprehension does not improve by simply reading more text (e.g., Pressley, Wharton-McDonald, Mistretta-Hampston, & Echevarria, 1998). Many students need to be explicitly taught strategies for understanding what they read (Gough & Hillinger, 1980; Wren, 2002). Pressley and colleagues (1998) stated that by learning and using even one strategy, an individual's comprehension can improve. Additionally, by learning and using multiple strategies, an individual's comprehension may improve even more significantly (Pressley et al., 1998). By having explicit strategies to use when they are reading, students are able to more accurately understand what they are reading when they are reading it. These strategic approaches offer students a way to systematically approach reading and subsequently comprehension oftentimes increases (e.g., Gough & Hillinger, 1980; Pressley et al., 1998; Wren, 2002).

In their quasi-experimental study Anderson and colleagues (1995) examined the differences in students' SAT comprehension scores between a treatment (strategy instruction) and control (current school practices) group. They found a statistically significant difference in scores from pre to post-test from the treatment group ($p < .05$, $d = 1.16$). Additionally, in their review of the research on literacy studies conducted with secondary struggling readers in grades 6-12 from 1994-2004, Edmonds and colleagues (2009) found interventions associated with strategy instruction and comprehension outcomes ($n = 13$) yielded an effect size of .89. That is, students in the treatment conditions scored, on average, more than two thirds of a standard deviation higher than students in the comparison conditions on measures of comprehension (Edmonds et al., 2009). These studies support the notion that struggling and proficient readers benefit from explicit cognitive strategy instruction targeted to increase comprehension (e.g., Fisher & Frey, 2011; Fisher et al., 2012).

Although researchers have not concretely agreed on which approach to reading strategy instruction produces the greatest student gains, they do contend that explicit instruction and continued use of these strategies (e.g., activating background knowledge, self-monitoring skills, text-based questioning) are essential for students to read, interpret, and critique a variety of texts (Brown, 2002; Guthrie, Wigfield, & Perencevich, 2004; Moje & Speyer, 2008; Taboada & Guthrie, 2006; Taboada & Rutherford, 2011, Vaughn et al., 2009). Research has shown that in order for students to become proficient content area readers, they should apply these cognitive reading strategies before, during, and after reading (Duke & Pearson, 2002; Paris & Oka, 1986; Vacca, 2002). When students engage in the use of these strategies they can make the literacy gains necessary to move from poor readers to systematic and strategic readers who can traverse a multitude of texts with a handful of research-based strategies.

Strategy use and student metacognition. Metacognition is the knowledge and control students have over their own thinking and learning activities and includes two broad categories of mental activities: self-appraised knowledge and self-management of one's cognition and thinking (Cross & Paris, 1988). Therefore, metacognition involves not only one's thoughts about their own thinking, but also knowing how to manage and strategically understand one's thinking. Put another way, metacognition includes knowledge about cognition and includes the variables that influence thinking and the ability to act accordingly (Flavell, 1979; Garner, 1987).

Although there are many approaches to metacognition, for the scope of this paper I am focusing on Paris and Lindauer's (1982) three category approach (i.e., evaluation, planning, regulation) to metacognition. In the context of reading, evaluation refers to analyses of task characteristics and personal abilities that affect comprehension; planning involves the selection of particular strategies to reach the goals that have been set or chosen; and regulation is the monitoring and redirection of one's activities during the course of reading to reach the desired goals (Paris & Lindauer, 1982). These three characteristics of metacognition can be related to students' use of strategies and translate into before (evaluation), during (planning), and after reading (regulation) strategies. More specifically, before a proficient reader approaches a text they evaluate the reading task, then they strategically read and make a plan when comprehension breaks down, and finally they regulate themselves throughout the reading process to ensure understanding. Thus, there is a connection between students' metacognition and their abilities to use reading strategies in aid in comprehension when reading.

Using a strategic approach to reading demonstrates an ability for a student to metacognitively think about what they are reading. Skilled readers are able to use metacognition when reading to: assess their own knowledge, monitor their comprehension, and implement fix-up strategies when their comprehension fails (Anderson, Hiebert, Scott, & Wilkinson, 1985; Garner, 1987). Although proficient readers are able to think about their own thinking when reading (i.e., be metacognitive) and fix their comprehension when it fails (i.e., implement reading strategies), struggling readers often lack this metacognitive awareness leading to an overall lack of reading proficiency (Anderson et al., 1985; Garner, 1987; Moore, Zabrucky, & Commander, 1997; Ofodu & Adedipe, 2011). Thus, if a reader is not aware of his/her own cognitive processes, and the strategies used to support these processes, than the reader will be less likely to notice when gaps in understanding occur while reading; therefore, the reader will be less likely to know how to use strategies to bridge these gaps in understanding, which ultimately leads to limited comprehension. Thus it becomes important for teachers to engage in explicit strategy instruction that demonstrates to students how to metacognitively think about their own understanding when reading.

Strategy instruction and student metacognition. Strategy instruction and metacognition go hand in hand because strategy instruction encourages students to think about their own mental processes and then execute specific strategies to help them interact and make meaning of text (Gooden, 2012). Strategy instruction challenges all students to be metacognitively aware and thinking about their reading as they read. In order to be metacognitively aware, an individual must be conscious of their cognitive processes while they are doing a given task, in this case reading (Flavell, 1976, 1979). Thus, teachers must engage in instruction that brings the invisible aspects of metacognition (i.e., one's way of thinking), visibly into a classroom.

According to Vacca (2002), effective literacy instruction begins with well-planned content lessons, explicit comprehension instruction in the content area, and sustained literacy instruction before, during, and after students engage with texts. When literacy instruction occurs before content area reading (e.g., having students do a book walk before reading), background knowledge is built. Using these strategies can aid in students' motivation and engagement with texts as new vocabulary and concepts are introduced. Additionally, these strategies help to develop students' metacognitive awareness of pre-reading strategies (Boulware-Gooden, Carreker, Thornhill, & Joshi, 2007; Guthrie et al., 2004; Paris & Oka, 1986; Pressley & Harris, 1990; Tang & Moore, 1992). After students have engaged in pre-reading, the teacher scaffolds literacy skills during the reading phase of instruction. In the post-reading phase, the teacher focuses students in critical thinking to make content area connections by applying newly learned knowledge through interpreting, summarizing, and critiquing the texts explored (Duffy, 2002; Pressley, 2002).

Additionally, in order for students to become more metacognitively aware of what strategies to use in what contexts, some literacy researchers recommend that strategy instruction should follow Pearson and Gallagher's (1983) Gradual Release of Responsibility (GRR) model of modeling, guided practice, and independent practice. Adolescent literacy strategy instruction using the GRR model involves the teacher modeling the use of reading strategies, guiding students through the correct use of these strategies, and then creating the opportunity for adolescents to independently practice the strategies all while using content area texts (Boulware-Gooden et al., 2007; Clark & Graves, 2005; Fisher & Frey, 2008; Pearson & Dole, 1987; Slater & Horstman, 2002). As students see the teacher model (vicarious experience), engage in guided practice (mastery experience), and participate in independent practice (mastery experience) they are able to utilize reading strategies across a variety of contexts in an effort to increase comprehension and reading proficiency. These literacy practices help students become more metacognitively aware of the various literacy skills needed to be successful in this literacy-based society (Cotterall & Murray, 2009; Mokhtari & Reichard, 2002; Pintrich & De Groot, 1990; Pressley & Gaskins, 2006; Yang, 1999; Zare-ee, 2007).

Reading motivation. Another area of interest within the content-area literacy community is the necessity for students to be engaged and motivated to read in order to increase student reading achievement. Reading motivation includes the goals and beliefs students hold in relation to reading, which influence their interactions and learning with texts (Guthrie & Wigfield, 1999). Students' reading motivation may be influenced by strategy instruction as increased strategy instruction has been found to lead to increased strategy use, which has been shown to increase students' motivation to read (Guthrie et al., 1996). This relationship occurs because skill and motivation have been found to be reciprocal; as students who increase their skills in reading may become more motivated, and conversely, as students become more motivated they may be inspired to continue to use their reading skills (Cambria & Guthrie, 2010; Guthrie & Wigfield, 1999). Content area researchers state that cognitive strategy instruction, explicit vocabulary instruction, time with text, and opportunities to engage in discussions about text meaning and interpretation, can lead to student motivation and engagement with texts and an overall building of knowledge through reading (Brozo & Simpson, 2006; Pressley, 2006; Taboada & Guthrie, 2006; Vaughn et al., 2009). Studies have shown that as students become more familiar with content area reading strategies their comprehension can improve, which can increase their motivation to read, and achievement may subsequently increase (see Irvin et al., 2007 for an example).

In their longitudinal study of instruction, reading motivation, and achievement, Guthrie and Wigfield (2000) concluded that although instructional practices are important, they alone do not directly impact student outcomes. Rather students' level of engagement is the mediating factor between instruction and outcomes. That is, students' level of engagement increases or decreases the impact of instruction on student outcomes. Therefore, even with excellent literacy instruction at the secondary level, if students are not motivated to read, adolescent literacy achievement will not improve; thus, students must be motivated to read in order for instruction to impact student outcomes. Thus, student motivation to read is a key component to the successful integration of content-area reading instruction in the secondary classroom. In order to more accurately meet the needs of adolescent readers, secondary teachers must engage in cognitive strategy instruction while also providing students with a motivating environment with which to engage with texts in an effort to improve the reading proficiency of adolescent students in the United States (Cambria & Guthrie, 2010; Fisher et al., 2012; Guthrie & Wigfield, 1999).

Struggling readers. Struggling adolescent readers are likely to encompass three primary categories: a) students who have cognitive difficulties that influence their ability to understand text (Worthy & Invernizzi, 1995); b) students who may be able to perceive and articulate the words in a text, but cannot make meaning from these words in order to understand the text (Kim & Goetz, 1994); or c) students who do not know how to set goals or apply strategies to understand text when their comprehension fails (Gersten et al., 2001). Struggling readers often have low reading motivation due in part to their difficulties with comprehension when reading. Therefore, in order to increase literacy achievement among proficient and struggling readers, teachers must incorporate cognitive reading strategy instruction into the secondary classroom while creating an environment that fosters students' motivation to read.

Using explicit strategy instruction and the GRR model (Pearson & Gallagher, 1983), teachers can provide struggling readers with concrete strategies to improve comprehension and subsequent motivation to read. The use of cognitive reading strategy instruction cannot only increase student reading achievement, but can also influence students' motivation to read through the reciprocal nature of skills and motivation. That is, as students' ability to use reading strategies effectively increases, so does their comprehension, and thus their motivation for reading. Additionally, as the need to gain meaning from text increases drastically as students progress through school, knowing the correct application of cognitive reading strategies becomes necessary for adolescent readers; especially those who struggle with comprehension (Biancarosa & Snow, 2004; Perfetti, Landi, & Oakhill, 2005). Therefore, in order to improve adolescent reading achievement in the US, secondary teachers must use content-area reading strategies in a motivating and supporting environment in order to successfully reach both struggling and proficient readers.

Disciplinary literacy. Another approach to adolescent literacy is disciplinary literacy which focuses on each discipline having its own discourse with specialized language, texts, and literacy practices inherent to that discipline (O'Brien et al., 1995; Shanahan & Shanahan, 2008). That is, disciplinary literacy researchers articulate the need for discipline specific over general reading strategies, as articulated by content area researchers. Teachers should instruct students on how to utilize a variety of strategies within each discipline (e.g., building discipline specific vocabulary, posing discipline specific questions).

According to Moje (2007), using disciplinary literacy practices students will be able to develop deep content knowledge and literature habits within each of the specific disciplines. This is the same fundamental principal of content-area literacy; but, within a disciplinary literacy framework, strategies must be specific to each discipline and not generic strategies used across content areas. Researchers in disciplinary literacy believe that adolescents should learn to read like historians, scientists, mathematicians, and literary thinkers and employ the discipline specific strategies (i.e., using historical inquiry when engaging in history texts) necessary to complete this task (Buehl, 2009; McConachie & Petrosky, 2009; Moje, 2008; Monte-Sano, 2011; Shanahan & Shanahan, 2008). Thus, disciplinary literacy focuses on the acquisition of knowledge through discipline specific literacy instruction instead of the general content-area literacy instruction that content area researchers discuss.

Shanahan and Shanahan (2008) have taken disciplinary literacy research one step further and proposed a model where “disciplinary literacy is the most advanced level of literacy and utilizes skills specialized to particular disciplines” (p. 44). According to their model, there are three categories of literacy: basic literacy (e.g., decoding, knowledge of high-frequency words); intermediate literacy (e.g., general comprehension strategies, common word meanings, and fluency); and disciplinary literacy, which is the most advanced level of literacy, where the skills are specialized to a particular discipline (e.g., history, science, mathematics). In this model, content-area literacy falls under intermediate literacy as it uses general comprehension strategies and, according to Shanahan and Shanahan (2008), teachers should strive for discipline specific literacy or advanced literacy instruction.

Further, disciplinary literacy learning focuses on students' ability to construct meaning from a multitude of technical and highly discipline specific texts and to use critical literacy skills to construct meaning in each discipline (Moje, 2008; O'Brien et al., 1995; Shanahan & Shanahan, 2008). As texts get more complex, students must learn the various discipline specific strategies to use to construct and interpret highly specialized texts. Instead of using a set of generic content-area reading strategies (e.g., activating background knowledge) to improve the reading of adolescent students, disciplinary literacy seeks to emphasize that students must learn the ways of knowing, doing, and communicating within a discipline (Moje, 2008).

Disciplinary literacy researchers embrace a perspective that each discipline has its own set of inherent beliefs and pedagogy that dictate how members of that discipline interact and engage with text (Bain, 2006; Wineburg, 1991; Yore, Hand, & Prain, 2002). This aligns with the belief that for true disciplinary knowledge to occur students must "think like a [insert discipline]" and code-switch their literacy skills to be able to effectively comprehend and interpret texts in that discipline. Additionally, within this framework students are required to apply basic skills to comprehend, interpret, and apply disciplinary knowledge (Moje, Collazo, Carrillo, & Marx; 2001). Yet, for students to acquire these skills, literacy instruction must occur at the secondary level.

Content-area literacy and disciplinary literacy. As previously shown, empirical evidence has provided insights into how content-area reading instruction has positively impacted student achievement (e.g., Fisher et al., 2002) and their motivation to read (e.g., Guthrie & Wigfield, 1999). However, the empirical evidence for disciplinary reading strategy instruction is sparse. Monte-Sano and De La Paz have conducted studies that link disciplinary writing instruction in history with student writing gains (i.e., De La Paz & Felton, 2010; Monte-Sano & De La Paz, 2012; Monte-Sano, De La Paz, & Felton, 2014), but none of their studies have embedded reading instruction or examined reading outcomes in social studies. For the scope of this study, only content-area literacy research was consulted based on empirical evidence and the more intermediate nature of content-area literacy.

Teacher Beliefs Influencing Literacy Instruction

Although the research community is divided over the methods by which secondary teachers should infuse literacy instruction in their classrooms, numerous literacy researchers have prompted content area teachers to accept responsibility for literacy development (Alvermann, 2001; Bean, 2000; Biancarosa & Snow, 2006; Elkins & Luke, 1999; Gray, 1925, Herber, 1978; Kamil, 2003; Moore et al., 1983; Vacca, 2002). Content area teachers play a pivotal role in adolescent literacy success, as they are the ones who can connect students' ability levels and interests to literacy within their content area (Darling-Hammond, 1997; Freedman & Carver, 2007). Therefore, literacy instruction must occur at the secondary level in order for teachers to have the ability to positively influence students' reading achievement.

Teachers' beliefs play a critical role in promoting or inhibiting successful classroom practices. As Pajares (1992) articulated, beliefs are crucial in defining a task and choosing how to approach, plan, and make decisions about said task; thus, beliefs are critical to how one behaves and obtains information. For teachers, their beliefs dictate how they approach instruction, difficulties with students, and their classroom environment. In relation to content-area literacy instruction, many teachers believe that literacy instruction will detract from the content and will waste too much class time (Fisher & Ivey, 2005; Heller & Greenleaf, 2007). Moreover, such beliefs cause secondary teachers to struggle or ignore literacy instruction in their classrooms (Conley, 2009; Fisher & Ivey, 2005; Sturtevant & Linek, 2003).

Furthermore, research has shown that even when secondary teachers want to use content-area literacy instruction in their classroom, they lack the training and resources necessary to effectively deliver literacy instruction (Gee & Forrester, 1988; Hall, 2005; Heller & Greenleaf, 2007; Sturtevant, 1996). Teachers' lack of training can lead to a lack of confidence (i.e., a lower sense of efficacy) and increased frustration and discomfort with content-area literacy instruction (Greenleaf, Schoenbach, Cziko, & Mueller, 2001). In order to combat a low sense of efficacy, researchers and teacher educators must work to develop teachers' knowledge and skills related to literacy instruction in the secondary classrooms (Biancarosa & Snow, 2006; Fisher & Ivey, 2005; Guskey, 2000; Olson &

Truxaw, 2009). Thus, teacher self-efficacy for literacy instruction is an essential component to improved and effective reading instruction at the secondary level.

Teacher Self-Efficacy

The field of self-efficacy has two guiding conceptual frameworks (Rotter, 1966; Bandura, 1977) that have delineated how the construct has been defined and measured. From these two conceptual frameworks, teacher self-efficacy became a field in which researchers have been divided on which framework of self-efficacy is best when devising instrumentation. Although the field is split, after first describing each framework this section focused on Bandura's (1977, 1986) definition and conception of self-efficacy as a multi-faceted, context, domain, and task-specific construct. Therefore, only findings from studies centered on these tenets were included when describing how teacher self-efficacy influences instruction and student outcomes. Additionally, measures created around Bandura's (1977, 1986) framework were described in greater detail than those that used Rotter's (1966) framework.

Rotter's Locus of Control

Following the work of Lewin in 1949, Rotter (1966) stated that all humans have an internal versus external locus of control that depicts how actions are viewed. Rotter (1966) defined this framework stating:

> When a reinforcement is perceived by the subject as following some action of his own but not being entirely contingent upon his action, we have labeled this belief in *external control*. If the person perceives that the event is contingent upon his

> own behavior or his own relatively permanent characteristics, we have termed this a belief in *internal control* (p. 1).

Hence, one's locus of control is the perceived source of control over one's behavior that influences the way one views himself/herself and his/her opportunities. Therefore, using the locus of control framework, self-efficacy (i.e., beliefs about one's ability) is the internal control and outcome expectancy (i.e., results from performance) is the external control. This theory provided the conceptual framework for one line of inquiry (e.g., Riggs & Enochs, 1990; RAND studies) into teacher self-efficacy around teachers' locus of control.

Many of the instruments based on the locus of control framework incorrectly measure outcome expectancy as an aspect of self-efficacy. Chief among the conceptual difficulties in scales based on Rotter's (1966) locus of control framework is a focus on teachers' beliefs about their control of student outcomes (outcome expectancy) rather than a focus on the teachers' capabilities to effectively teach students (self-efficacy). However, this measurement of self-efficacy and outcome expectancy as a related construct is a misconceptualization of Bandura's (1977) notion that self-efficacy and outcome expectancy are clearly distinguishable constructs. This provided another rationale for using Bandura's conception of self-efficacy over Rotter's locus of control framework.

Bandura's Social Cognitive Framework

Another conceptual framework of self-efficacy focused on Bandura's (1986) social cognitive theory that describes one's actions through triadic reciprocal determinism, or the idea that personal (P; e.g., cognition), behavioral (B), and environmental (E) factors influence one another in a reciprocal fashion. Bandura (1989) discussed this reciprocation as:

> The P ↔ B of reciprocal causation reflects the interaction between thought, affect and action. The E ↔P segment of reciprocal causation is concerned with the interactive relation between personal characteristics and environmental influences. The B ↔ E segment of reciprocal causation in the triadic system represents the two-way influence between behavior and the environment (pp. 3-4).

According to social cognitive theory, what people think, believe, and feel, impacts how they behave, and these thoughts, beliefs, and feelings are developed and changed through social persuasions which can be altered by the environment. This theory provided the conceptual framework for the other line of teacher self-efficacy research (e.g., Bandura, 1997; Tschannen-Moran & Woolfolk Hoy, 2001). Furthermore, Bandura (1977, 1986) articulated that self-efficacy was a multidimensional construct encompassed by context, domain, and task-specific demands and therefore all instrumentation should reflect this specificity of measurement.

The need for more specific measurement arises with the desire to predict student outcomes from teachers' sense of efficacy. As Pajares (1996, 1997) argued, self-efficacy instruments lose predictability when they measure the construct in a generalized way. Therefore, in order to correlate and attempt to predict student outcomes from teacher self-efficacy, both constructs must be measured in a context, domain, and task-specific manner. Additionally, within the social cognitive framework, self-efficacy develops from four sources of influence: mastery experiences (the first-person experiences one has), vicarious experiences (the social models one sees), verbal persuasions (the social cues one hears), and physiological state (the emotional state one is in; Bandura, 1977). These forms of influence work in tandem to formulate one's perceived conception of self-efficacy. Moreover, Bandura (1997) postulated that mastery experiences (i.e., the actual first-person experiences one has) influenced one's self-efficacy more than the other forms of influence.

When discussing the need for effective literacy instruction in the secondary content areas, it becomes pertinent to focus efforts into examining the perceived self-efficacy of secondary content area teachers in this type of instruction. In congruence with Bandura (1986) and Pajares (1996), in order to effectively predict student outcomes from teachers' sense of efficacy, both constructs must be measured on the same level of specificity. Therefore, in accordance with the adolescent literacy literature and the need for more effective literacy instruction at the secondary level, a new teacher self-efficacy instrument should be created based on Bandura's conception of self-efficacy. This new measure should address teacher self-efficacy from a context (i.e., secondary), domain (i.e., content areas), and task-specific (i.e., content –area reading instruction) manner and account for the mastery experiences (i.e., in-service teachers) of its target population.

Context specificity. As Bandura (1986, 1997) articulated, self-efficacy is not a context-less global disposition and therefore should not be measured by global omnibus instruments. Therefore, measures should be context-specific as context influences one's sense of efficacy. When discussing teacher self-efficacy it is important that studies focus around a particular group/level of teaching (e.g., secondary teachers) as teaching varies greatly from one level to another. Elementary and secondary schools provide vastly different contexts for teachers (Di Fabio & Palazzeschi, 2008) and so teachers' sense of efficacy may be influenced by their specific teaching context. For example, in the US many elementary teachers teach a variety of subjects (e.g., math and science) to one or two groups of students, whereas secondary teachers teach one content area (e.g. science) to multiple groups of students throughout the day. In line with Bandura's conception of self-efficacy as context-specific, elementary and secondary teachers would report self-efficacy beliefs based upon the context with which they are teaching.

Although Bandura defined self-efficacy as a context-specific construct, much of the current self-efficacy research has focused on elementary teachers specifically, or used mixed samples of teachers; leading to an underrepresentation of secondary specific studies. Klassen and colleagues (2011) found that of the 218 studies they examined 24% (n = 52) were focused on elementary teachers, with 15% (n = 33) concentrated on high school teachers, and an additional 25% (n = 54) where multiple teaching levels were examined. Additionally, many of the studies that encompassed a solely secondary sample of teachers, utilized general measures of teacher self-efficacy (e.g., Cantrell & Hughes, 2008; Table 1) or a scale was developed solely for use in the study without replication and further validation (e.g., McNeill, Pimental, & Strauss, 2013; Table 1).

Table 1

Self-Efficacy Studies with Samples of only Secondary Teachers

Authors	**Participants**	**Methodology**	**Self-Efficacy Measure**
Barros, Labruru, & Da Silva (2010)	136 Brazilian physics teachers	Quantitative	Created measure, 18 items, 2 factors: (1) General Efficacy Belief in Physics Teaching (2) Personal Efficacy Belief in Physics Teachers
Behar-Horenstein, Pajares, & George (1996)	14 secondary teachers	Mixed Methods	Interviews
Betoret (2006)	247 Spanish secondary teachers	Quantitative	Created measure, 7 items, 1 factor: (1) Self-efficacy
Cantrell, Burns, & Callaway (2009)	28 sixth and ninth grade teachers	Qualitative	Interviews
Cantrell & Callaway (2008)	16 sixth and ninth grade teachers	Qualitative	Interviews
Cantrell & Hughes (2008)	22 sixth and ninth grade teachers	Mixed Methods	Created measure, 65 items, 2 factors; adapted items for reading from Woolfolk and Hoy (1990), Hoy and Woolfolk (1993), and Gibson and Dembo (1984) with 2 factors: (1) General Teaching Efficacy (GTE) (2) Personal Teaching Efficacy (PTE)

Chacon (2005)	104 Venezuelan high school English as a Foreign Language teachers	Mixed Methods	Created measure-English Teachers' Sense of Efficacy Scale (ETSES), 39 items, 5 factors; factors 1-3 adapted items from TSES, Tschannen-Moran and Woolfolk Hoy (2001): (1) teachers' perceived efficacy for engaging students in learning EFL (2) teachers' perceived efficacy for managing EFL classes (3) teachers' perceived efficacy for implementing instructional strategies to teach EFL (4) teachers' self-reported English proficiency (5) teachers' self-reported pedagogical strategies to teach English
Caprara, Barbaranelli, Borgogni, & Steca (2003)	2688 Italian middle school teachers	Quantitative	Created measure, 12 items, 1 factor: (1) Perceived self-efficacy
Caprara, Barbanelli, Steca, & Malone (2006)	2184 Italian middle school teachers	Quantitative	Created measure, 7 items, 1 factor: (1) Teachers' self-efficacy beliefs
Chan (2004)	158 secondary school teachers in China	Quantitative	Wegner–Schwarzer–Jerusalem 10-item general self-efficacy scale (GSE; Schwarzer, 1993); Schwarzer–Wegner 10-item self-efficacy toward helping scale (SETH; Schwarzer, 1993)
Deemer (2004)	99 high school science teachers	Quantitative	Created measure-Personal Teaching Efficacy (TEFF), 8 items, 1 factor; adapted from TES by Gibson and Dembo (1984): (1) Personal Teaching Efficacy

Di Fabio & Palazzeschi (2008)	169 Italian high school teachers	Quantitative	TSES, 24 items, 3 factors; Tschannen-Moran & Woolfolk Hoy (2001): (1) self-efficacy for instructional strategies (2) self-efficacy for classroom management (3) self-efficacy for student engagement
Ebrahimi & Moafian (2012)	164 Iranian high school English as a Foreign Language teachers	Quantitative	TSES, 24 items, 3 factors; Tschannen-Moran & Woolfolk Hoy (2001): (1) self-efficacy for instructional strategies (2) self-efficacy for classroom management (3) self-efficacy for student engagement
Evers, Brouwers, & Tomic (2002)	490 Dutch teachers in the upper grades of schools of general vocational education	Quantitative	Created measure, 13 items, 3 factors: (1) to guide groups of students using the principle of differentiation (2) to involve pupils in tasks (3) the use of innovative educational practices
Ghaith & Yaghi (1997)	25 secondary teachers	Quantitative	TES, 16 items, 2 factors; Gibson and Dembo (1984): (1) GTE (2) PTE
Holzberger, Philipp, & Kunter (2013)	155 German secondary math teachers	Quantitative	Teacher Self-Efficacy Scale, 4 items, by Schwarzer, Schmitz, and Daytner (1999)
Holzberger, Philipp, & Kunter (2014)	155 German secondary mathematics teachers	Quantitative	Teacher Self-Efficacy Scale, 5 items, by Schwarzer, Schmitz, and Daytner (1999)

Kurz & Knight (2004)	113 high school teachers	Quantitative	TES, 16 items, 2 factors; Gibson and Dembo (1984): (1) GTE (2) PTE
Lee, Dedrick, & Smith (1991)	8,488 secondary teachers in public/Catholic schools	Quantitative	Created measure, 4 items, 1 factor: (1) Teaching Self-Efficacy
McNeill, Pimental, & Strauss (2013)	22 secondary ecology teachers	Quantitative	Created measure, 6 items, 1 factor: (1) Comfortable teaching students to design and conduct investigations
Raudenbush, Rowen, & Cheong (1992)	315 secondary teachers	Quantitative	One item: “To what extent do you feel successful in providing the kind of education you would like to provide for the students in this class” with a 4-point scale ranging from “no successful” to “highly successful”
Ross (1992)	18 Canadian secondary history teachers	Quantitative	TES, 16 items, 2 factors; Gibson and Dembo (1984): (1) GTE (2) PTE
Ross, Cousins, & Gadalla (1996)	52 secondary teachers	Quantitative	RAND, 1 item, Armor et al., (1976): “If I try really hard I can get through to even the most difficult or unmotivated students in this class”
Van Uden, Ritzen, & Pieters (2013)	195 Dutch teachers in prevocational and vocational education	Quantitative	Bandura’s (2006) questionnaire, 14 items, 3 factors: (1) instructional self-efficacy, (2) disciplinary self-efficacy, and (3) efficacy to create a positive school climate

Warren & Payne (1997)	82 eighth grade teachers	Quantitative	TES, 30 items, 2 factors; Gibson and Dembo (1984): (1) GTE (2) PTE

Note. Data as reported in research article.

Thus, there is a gap in the literature in the study of secondary teachers' sense of efficacy using validated and reliable secondary-specific (context) measures. If primary and secondary teachers have different teaching contexts, it becomes necessary to devise self-efficacy instruments to measures these beliefs contextually. Additionally, context-specific self-efficacy measures should be created that reflect the unique teaching demands of the target population (i.e., elementary or secondary teachers), instead of global measures used across both groups.

Domain specificity. Aligning with Bandura's (1986) definition, perceived self-efficacy is domain-specific, thus one may have high self-efficacy in one activity domain but not necessarily high self-efficacy in another. Thus, self-efficacy measures should be created that focus on assessing perceived self-efficacy in a particular domain(s) to adhere to Bandura's (1986) conception of self-efficacy. In relation to teaching, the domain (i.e., content or subject area) may affect a teachers' sense of efficacy and this level of self-efficacy is not necessarily the same across subject areas. Further, Bandura (1997) and Pajares (1997) stated that teacher efficacy scales should be linked to specific domains in order to properly assess how the teachers' sense of efficacy contributed to students' academic achievement. Domain-specific measurement is essential as teachers' may feel efficacious in certain domains but not in others, an aspect of self-efficacy that is not captured by generalized measures of teacher self-efficacy.

Despite both Bandura (1997) and Pajares (1996) delineating the need for domain-specific teacher self-efficacy measurement, Klassen and colleagues (2011) found in their review of the teacher self-efficacy research that 60% (*n* = 130) of studies measured perceived sense of efficacy using a general measure. In addition, the most commonly used teacher self-efficacy measures (i.e., TES and TSES) are both global measures of teacher self-efficacy that suffer from generally worded items asking teachers to assess their efficacy for teaching, without defining a particular domain, or subject area. Even though the TSES examined teacher self-efficacy in three task-specific areas (i.e., self-efficacy for classroom management, instructional practices, and student engagement), it still suffers from a non-specified teaching domain. For example, one item from the TSES states, "How well can you respond to difficult questions from your students?" (Tschannen-Moran & Woolfolk Hoy, 2001). Without a domain specified the teacher must answer the question from a general teaching standpoint.

Additionally, it becomes difficult to make interpretations about teachers' sense of efficacy when an instrument measures global self-efficacy because there is no connection to a specific domain; therefore, it is impossible to determine the specific domain from which the teacher was responding (Wheatley, 2005). For example, a secondary English teacher may teach various courses within the English domain (e.g., English and journalism) and have difference perceived self-efficacy across courses; however, with a general measure of self-efficacy the nuanced differences become lost, as the teacher must respond to his/her overall teaching self-efficacy. When teacher self-efficacy instruments measures the construct in a globalized way, the domain-specific differences in perceived self-efficacy are not distinguishable and thus disappear.

Even though the most commonly used teacher self-efficacy scales (i.e., the TES and TSES) are general in nature, the creation of domain-specific measures has risen in the past three decades. These domain-specific instruments are mostly focused on the domain of science and were written around Rotter's (1966) locus of control framework. In their review Klassen and colleagues (2011) found that 15% (n = 32) of domain-specific instruments are focused on the domain of science, with only 2% (n = 4) focusing on language and literacy. In addition, within Bandura's (1977) social cognitive framework only four domain-specific teacher self-efficacy measures were found in the current review, and none focused around the perceived self-efficacy of secondary (context) content area (domain) teachers.

Task specificity. In addition to being context and domain-specific, perceived self-efficacy is also task-specific. Self-efficacy beliefs can change from task-to-task within a specific domain and context, as some tasks are inherently more difficult than others (Bandura, 1986; Pajares, 1996). Therefore, if the task is easy to perform without inherent difficulties, people would have high perceived self-efficacy; however, if the task is more demanding and has inherent difficulties associated with it, people would have different levels of perceived self-efficacy (Bandura, 1997). As a task-specific construct, self-efficacy measures should be designed to assess teachers' sense of efficacy in executing a specific task (e.g., content-area literacy instruction).

Bandura (1986, 1997) noted that in order to match student outcomes to teacher self-efficacy, teacher self-efficacy must be measured in a task-specific manner; otherwise, one is trying to predict student outcomes within a specific task with generalized teacher self-efficacy which may not be causally related. For example, trying to correlate students' math achievement with teachers' general teaching efficacy is not possible, as these variables are not both being measured task specifically, decreasing the ability to draw causal or predictive outcomes. Thus, both Bandura (1997) and Pajares (1996) articulated the desire for researchers to design task (e.g., content area literacy) specific measures of teacher self-efficacy in order to increase the predictive qualities of these instruments and to relate beliefs to outcomes.

Further, content-area reading instruction is a task that is not addressed in the current teacher self-efficacy instrumentation. As content-area reading instruction is a multi-faceted task, according to the content-area literacy literature, three specific tasks within reading instruction should be addressed in creation of a new measure: using content-area reading strategy instruction (e.g., a major component of content area literacy; Vacca, 2002), motivating students to read (e.g., a critical component in self-efficacy research as well as a driving force behind content area strategy instruction; Guthrie & Wigfield, 2000), and teaching struggling readers (e.g., 70% of the adolescent population is considered a struggling reader; Biancarosa & Snow, 2006). By defining these three areas within the content-area literacy framework, the measure can more accurately address teacher self-efficacy in a much more task-specific manner.

Target population and mastery experiences. As hypothesized by Bandura (1986), mastery experiences may influence one’s self-efficacy more so than any other form of persuasion. That is, the first-person experiences a person has will influence that person’s self-efficacy above and beyond the other three sources of influence. When thinking of teaching, a teacher’s prior experiences in the classroom will have the most positive or negative impact on his/her self-efficacy above and beyond anything else. For example, a teacher who has had great success will have a higher sense of efficacy, whereas a teacher who has accrued many failures may have a lower sense of efficacy. In their review of teacher self-efficacy literature, Klassen and colleagues (2011) found that 29% (n = 64) of studies focused primarily on pre-service teachers, whereas 25% (n = 54) utilized multiple levels of pre and in-service teachers. The issue with using a mixed sampling of teachers (i.e., pre and in-service) is Bandura's (1977, 1986, 1997) hypothesis that mastery experiences influence one's self-efficacy more than all other forms of persuasions. Therefore, teachers at different stages in their career may respond to self-efficacy measures differently, as was shown by Duffin, French, and Patrick (2012), Fives and Buehl (2010), and Henson (2002) all of whom found that the factor structure of the TSES was not the same for pre-service and in-service teachers.

Accordingly, the need arises for instrumentation specifically designed for the target population (i.e., pre service or in service teachers) as these two groups fundamentally have different mastery experiences and different assessments of their teaching self-efficacy. When examining teacher self-efficacy it is important to only focus on one group (i.e., pre or in service teachers) at a time as the mastery experiences between the two groups are different, and how they respond to the measure will be different. Research has shown that pre and in-service teachers' self-efficacy should be measured using different instrumentation (e.g., Henson 2002), yet many instruments and studies do not make this distinction.

Whereas there are a multitude of instruments that were designed to capture the self-efficacy of pre-service teachers (e.g., Haverback, 2007; Haverback & Perault, 2011;

Riggs & Enochs, 1990) all of which were validated and made reliable by utilizing samples of only pre-service teachers; this same trend does not continue for in-service teachers. Moreover, some of the most commonly used teacher self-efficacy instruments were validated using mixed samples of pre and in-service teachers (i.e. TES, Gibson & Dembo, 1984; TSES, Tschannen-Moran & Woolfolk Hoy, 2001). Since research has shown that pre and in-service teachers (e.g., Fives & Buehl, 2010) may respond to the same self-efficacy measure differently, it becomes highly problematic to use a mixed sample of pre and in-service teachers when attempting to provide validity and reliability.

Teacher Self-Efficacy and Instruction

Teacher sense of efficacy is a critical component to successful classrooms and ranks as a significant teacher characteristic associated with instructional quality/support (Guo, Dynia, Pelatti, & Justice, 2014; Holzberger, Philipp, & Kunter, 2013). For example, for in-service teachers, high teacher self-efficacy has been found to correlate with higher end of the year goals for students, positive teacher practices and policies used in the classroom, and innovative classroom techniques, (Caprara, Barbaranelli, Steca, & Malone, 2006; Evers, Browers, & Tomic, 2002; Guo, Connor, Roehrig, & Morrison, 2012), and has been shown to positively influence teachers' beliefs about teaching and instructional behaviors (Skaalvik & Skaalvik, 2007; Tschannen-Moran & Woolfolk Hoy, 2001).

Teaching self-efficacy beliefs also alter how much effort a teacher puts forth, how long he/she will persevere when confronting obstacles, and how resilient he/she is in the face of challenges (Hoy & Spero, 2005; Pajares, 1996; Tschannen-Moran et al., 1998).

Additionally, teachers with high self-efficacy beliefs were more likely to: implement curriculum innovations, use successful classroom management approaches and teaching methods that encourage student autonomy, manage classroom problems effectively, keep students on task, take responsibility for students with special needs, and have fruitful collaborative relationships with colleagues and parents that contributed to sustained work satisfaction and higher teacher retention (Betoret, 2006; Caprara, et al., 2006; Chan, 2006).

Teacher Self-Efficacy and Student Achievement

Teachers are one of the most important factors that impacts student learning (Bandura, 1993; Sanders et al., 1997) as they have the greatest control of the classroom environment through the practices they implement in their classrooms, and their instruction impacts student achievement (Guo et al., 2010; Ross, 1992; Sanders Wright & Horn, 1997). Teacher self-efficacy has been linked to higher: student achievement (Bruce, Esmonde, Ross, Dookie, & Beatty, 2010; Caprara et al., 2006; Palardy & Rumberger, 2008), student literacy gains (Guo et al., 2012; Guo et al., 2010; Guo, Justice, Sawyer, & Tompkins, 2011), student motivation (Bandura, 1993; Guo et al., 2011; Martin, Sass, & Schmitt, 2012), and students' perceptions of their teacher (Holzberger et al., 2013, 2014).

Furthermore, Bruce and colleagues (2010) argued that teacher self-efficacy alone may only marginally aid in student achievement; however, teacher goal setting, persistence in helping low achieving students, and use of challenging instruction strategies mediate the relationship between teacher self-efficacy and student achievement.

Additionally, they articulate how teacher's self-efficacy and professional actions correspond; therefore, teachers with more effective and rigorous instruction have higher self-efficacy and subsequent higher student achievement than their lower self-efficacy counterparts.

Measurement of Self-Efficacy

Current instrumentation in the field of teacher self-efficacy is designed surrounding either Rotter or Bandura's framework. Measures embody one of the theoretical frameworks from which the construct is defined and items are subsequently written. Historically, self-efficacy measurement was centered on Rotter's (1966) locus of control framework; however, as the field progressed more measures that embodied Bandura's (1977) social cognitive framework were created. Following the historical progression of teacher self-efficacy research, first measures based off Rotter's framework were introduced and then those devised off of Bandura's framework were discussed.

Table 2

Factor Loadings, Reliabilities, and Variance for Teacher Self-Efficacy Instruments

Instrument	**Items/Scale**	**Factors**	**Factor Reliabilities**	**Factor Loadings**	**Variance Explained**
Measures based on Locus of Control Theory					
RAND (Armor et al., 1976)	2 items; 5-point Likert scale	General Teaching Efficacy (GTE) Personal Teaching Efficacy (PTE) Combined for Teaching Efficacy (TE)			
Teacher Locus of Control (TLC; Rose & Medway, 1981)	28 items; 2-point scale	Factor 1: I- (external) Factor 2: I+ (internal)	I- = .81 I+ = .71	I-: .30 - .68 I+: .17 - .61	
Responsibility for Student Achievement (RSA; Guskey, 1981)	30 items; 100 point distribution	R+ (student success) R- (student failures)	R+ = .72 R- = .78	R+: .28-.72 R-: .32-.54	60.9%
Teacher Efficacy Scale (TES; Gibson & Dembo, 1984)	30 items; 6-point Likert scale	Personal Teaching Efficacy (PTE) General Teaching Efficacy (GTE)	PTE = .78 GTE = .75	PTE: .46-.61 GTE: .45-.65	28.8%
Adapted Teacher Efficacy Scale (TES; Soodak & Podell, 1996)	21 items; 6-point Likert scale	Personal Efficacy (PE) Outcome Efficacy (OE) Teaching Efficacy (TE)	PE = .80 OE = .73 TE = .70	PE: .40-.68 OE: .54-.76 TE: .36-.66	30.7%

Science Teacher Efficacy Belief Instrument (STEBI; Riggs & Enochs, 1990)	25 items; 5-point Likert scale	Personal Science Teaching Efficacy (PSTE) Science Teaching Outcome Expectancy (STOE)	Scale A: PSTE = .92 STOE = .77 Scale B: PSTE = .90 STOE = .76	PSTE: .47 - .81 STOE: .27-.76	
Microcomputer Utilization in Teaching Efficacy Beliefs Instrument (MUTEBI; Enochs, Riggs, & Ellis, 1993)	21 items; 5-point Likert scale	Outcome Expectancy (OE) Personal Self Efficacy (SE)	OE = .78 SE = .91	OE: .36-.63 SE: .46-.78	40.9%
Mathematics Teaching Efficacy Beliefs Instrument (MTEI; Enochs, Smith, & Huinker, 2000)	21 items; 5-point Likert scale	Personal Mathematics Teaching Efficacy (PMTE) Mathematics Teaching Outcome Expectancy (MTOE)	PMTE = .88 MTOE = .77	PMTE: .36-.65 MTOE: .42-.53	
Reading Teaching Efficacy Instrument (RTEI; Szabo & Mokhtari, 2004)	16 items; 5-point Likert scale	Self-Efficacy Scale (RTSE) Outcome Expectancy Scale (RTOE)	RTSE = .61 RTOE = .80	RTSE: .42-.68 RTOE: .45-.59	
Measures based on Social Cognitive Theory					
Exceptional Children who are English Learners (EXCEL) Teacher Inventory (Paneque & Barbetta, 2006)	20 items; 9-point scale		.94		

Ohio State Teaching Efficacy Scale (OSTES)/ Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Woolfolk Hoy, 2001)	24 items; 9-point Likert scale (Long)	Student Engagement (SE) Instructional Strategies (IS) Classroom Management (CM)	SE = .87 IS = .91 CM = .90	SE: .47-.75 IS: .59-.72 CM: .50-.78	58.5%
	12 items; 9-point scale (Short)	Student Engagement (SE) Instructional Strategies (IS) Classroom Management (CM)	SE = .81 IS = .86 CM = .86	SE: .62-.75 IS: .63-.73 CM: .61-.83	69.1%
Examining Teachers' Sense of Efficacy Scale (TSES; Fives & Buehl, 2010)	24 items; Long Form	1. Teacher Self-Efficacy (TSE) 2. 3 Factors above: SE, IS, CM	1. .95 2. .93	1. .61 or higher 2. .45-.83	1. 47.98% 2. 57.09%
1=Pre-service 2=In-service	12 items; Short Form	1. Teacher Self-Efficacy 2. 3 Factors above: SE, IS, CM	1. .92 2. .86	1. .60 or higher 2. .51-.79	1. 52.88% 2. 64.99%
English Language Teacher Efficacy Instrument (ELTEI; Akbari & Tavassoli, 2014)	32 items; 5-point scale	Classroom Management and Remedial Action (CMRA) Classroom Assessment and Materials Selection (CAMS) Skill and Proficiency Adjustment (SPA) Teaching and Correcting Language Components (TCLC) Age Adjustment (AA) Social Adaptation (SA) Core Efficacy (CA)		CMRA: .31-.66 CAMS: .31-.67 SPA: .35-.77 TCLC: .37-.76 AA: .49-.76 SA: .68-.73 CA: .47-.71	58.2%

Teachers' Sense of Efficacy for Literacy Instruction (TSELI; Tschannen-Moran & Johnson, 2011)	34 items; 9-point scale	Teachers' sense of efficacy in literacy instruction	.96	.63-.83	55%
Teachers' Efficacy Beliefs System-Self Form (TEBS-Self; Dellinger, Bobbett, Olivier, & Ellett, 2008)	30 items; 4-point scale (Olivier, 2000)	Communication/Clarification (CC) Management/Climate (MC) Accommodation of Individual Differences (AID) Motivation of Students (MS) Higher Order Thinking Skills (HOTS)	CC/MFL = .86-.87 MC/CM/PCC = .85-.86 PAI/AID = .85-.87 MS = .78	CC:.52-.72 MC: .58-.73 AID: .49-.81 MS: .56-.63 HOTS: .50-.80	61.9%
	30 items; 4-point scale (Bobbett, 2001)	Classroom Management (CM) Communication/Clarification (CC) Planning and Accommodating Individual Differences (PAI) Instilling Higher Order Thinking Skills (HOTS)	MLR = .80 HOTS = .85-.86	CM: .52-.71 CC: .53-.77 PAI: .55-.79 HOTS: .62-.77	58%
	30 items; 4-point scale (Dellinger, 2001)	Accommodation of Individual Differences (AID) Maintaining Positive Classroom Climate (PCC) Monitoring and Feedback for Learning (MFL) Managing Learner Routines (MLR)		AID: .61-.83 PCC: .73-.92 MFL: .55-.83 MLR: .51-.66	61.4%

Note. Empty cells represent information not included in article.

Locus of Control Framework

Research surrounding teachers' sense of efficacy began with the RAND studies of reading instruction in 1976. The RAND study assessed the extent to which teachers believed they could motivate their students outside of the control of external forces (Armor et al., 1976). In relation to teaching, Tschannen-Moran and colleagues (1998) explained that teachers who believed the environment was the main influence on their students' achievement had an external locus of control, and teachers who believed that they could control student achievement would have an internal locus of control.

The RAND scale was a two-item instrument that measured teachers' General Teaching Efficacy (GTE) and Personal Teaching Efficacy (PTE; Armor et al., 1976) with the idea of capturing how teachers felt about their teaching. The GTE measured teachers' external locus of control, the PTE measured their internal locus of control, and the sum of these scores was their Teaching Efficacy (TE). Armor and colleagues (1976) found that teachers with higher levels of self-efficacy believed they could influence student achievement and motivation. This study helped create the construct of teacher self-efficacy and laid the groundwork for all subsequent studies.

As researchers became more intrigued by the findings of the RAND studies, reliability concerns arose because the measure was only two items, leading to the development of longer and more encompassing scales. This led to the current multi-dimensional aspects of teacher self-efficacy scales. Two such scales based primarily off of the RAND study were the 28-item Teacher Locus of Control (TLC) scale (Rose & Medway, 1981) and the 30-item Responsibility for Student Achievement (RSA) scale

(Guskey, 1981); both based on Rotter's locus of control theory and designed to measure teachers' sense of efficacy in relation to GTE versus PTE (i.e. external versus internal locus of control). Additionally, Guskey (1981) introduced the idea that other teacher variables (e.g., gender, years of experience, grade level taught) may predict teacher sense of efficacy, which influenced much of the current research on teachers' sense of efficacy. The RAND scale began the research goal of capturing teachers' self-efficacy using instrumentation, and the TLC and RSA offered researchers multi-dimensional scales from which to measure the construct. Many of the current teacher self-efficacy measures were created based upon these initial three scales.

Teacher self-efficacy and outcome expectancy. As teacher self-efficacy research expanded, measures were created that captured another facet of teacher self-efficacy: outcome expectancy. As defined by Bandura (1986), "Perceived self-efficacy is a judgment of one's capability to accomplish a given level of performance, whereas as outcome expectation is a judgment of the likely consequences such behavior will produce" (p. 391). Although Bandura clearly differentiates between self-efficacy and outcome expectancy, many researchers have not upheld this distinction and have defined both constructs as a single entity. For example, in their seminal piece Gibson and Dembo (1984) articulated self-efficacy as a teacher's perceived ability to produce positive student gains, and outcome expectancies as to how much an environment could be controlled. Although the authors claim to align themselves with Bandura's (1986) definition of outcome expectancies, their definition is more aligned with Rotter's (1966) locus of control and whether teachers believe they can control environmental factors.

Additionally, Gibson and Dembo's definition is more related to teachers' general beliefs about teaching and not necessarily their beliefs about their effectiveness as teachers. Finally, their definition of self-efficacy deals more with teachers impacting school outcomes or change than teachers' beliefs in their own inherent abilities; again, dealing with actions outside of the teacher (external) rather than inside the teacher (internal). From this body of work came one of the most utilized and highly scrutinized measures of teacher self-efficacy the TES.

General measures based on locus of control. Gibson and Dembo (1984) created a 30-item scale to measure teacher's general self-efficacy and outcome expectancy. Based on 16 of the original 30 items, the authors found a two factor structure that they labeled: personal teaching efficacy and teaching efficacy (Gibson & Dembo, 1984). Although this is one of the most commonly used teacher self-efficacy measures, it is general in nature and therefore does not function well when trying to link self-efficacy and outcomes. Additionally, Gibson and Dembo's (1984) instrument was more closely aligned with Rotter's locus of control than with Bandura's social cognitive theory. This was demonstrated in the work of Guskey and Passaro (1994) who found the two factor structure to be internal versus external control, rather than personal and general teaching efficacy as stated by Gibson and Dembo (1984).

Conceptual issues with TES. In an attempt to study the factor structure of the TES, Guskey and Passaro (1994) conducted a factor analysis. Before they began the study Guskey and Passaro (1994) did a "close inspection" of the items and determined that all the items that loaded on the personal teaching efficacy factor were positively worded, included the word *I*, and had an internal locus (i.e. “I can”) and all the items that loaded on the teaching efficacy factor were negatively worded, included the word *teachers*, and had an external locus (i.e. “teachers cannot). Thus, it was unclear whether the factors were truly personal versus teaching efficacy, or rather internal versus external locus of control (Guskey & Passaro, 1994). In order to examine this, the authors sought to change the wording of items in the original TES scale to determine if the factors were more aligned with Gibson and Dembo (1984) reported personal teaching efficacy and teaching efficacy or with Rotter’s (1966) internal versus external locus of control.

When re-wording items Guskey and Passaro (1994) changed seven of the twelve personal efficacy items from using the word *I* to using the term *the teacher*, and changed four of the nine teaching efficacy items from using the word *the teacher* to the term *I*. They also added the two RAND items, and three items that Woolfolk and Hoy (1990) found to be significant when they used the TES. The result was a 21-item adapted TES instrument. Next, the authors conducted a factor analysis to generate the two-factor solution Gibson and Dembo (1984) discovered. Their data revealed that 11 items loaded on factor 1 and were negative and external, and the other 10 items loaded on factor 2 and were positive and internal (Guskey & Passaro, 1994).

Even though Guskey and Passaro (1994) found a two-factor solution, the items did not differentiate the factors as they had in the Gibson and Dembo (1984) study. Instead, both personal and teacher efficacy items loaded on factor 1 and factor 2; suggesting the factor patterns that Gibson and Dembo (1984) described as personal teaching efficacy and teaching efficacy did not hold up when the terms *I* and *the teacher* were changed. Instead, Guskey and Passaro (1994) found the difference to be an internal versus external distinction not the original personal versus general teaching factors. Additionally, they described the distinction between the two factors as the internal factor represented perceptions of personal power and the external factor represented perceptions of power outside of the classroom (Guskey & Passaro, 1994). Therefore, Guskey and Passaro (1994) demonstrated that the factor structure of the TES was aligned with Rotter's locus of control and not Bandura's social cognitive theory.

Further, Henson (2002) explained that the TES was based upon the RAND measure which was based upon the locus of control theory; leading to some items more closely resembling the external versus internal locus of control than Bandura's social cognitive theory. Bandura operationalized self-efficacy and outcome expectance as two uncorrelated concepts, therefore, the difficulty with the TES is the blending of both constructs. For the purposes of this study, all instruments that were modeled after the TES were discussed briefly, as these inherently followed this same pattern and align with Rotter's locus of control theory.

Although the aforementioned studies provided rationale for validity concerns with the TES (e.g. the wording of items, the two-factor structure is not all encompassing) many studies continued to assess teacher self-efficacy using the TES in a variety of forms as shown in the review of literature stating one-third (n = 72) of the teacher self-efficacy studies from 1998-2009 (n = 218) used some version of the TES (Klassen et al., 2011). Hoy and Woolfolk (1993) noted that one of the main issues with the TES is "that the first dimension of efficacy does not represent an outcome expectation as defined by Bandura" but rather a more “general belief about the power of teaching” (p. 357). Additional issues arose when Coladarci and Fink (1995) found weak evidence for discriminant validity of PTE and GTE scores, further questioning the validity of the TES measure. Also, when factor reliabilities were examined, many studies (e.g., Guskey & Passaro, 1994) found weak reliability in the data produced (Table 2), especially with the GTE factor (Coladarci & Fink, 1995; Henson, Kogan, Vacha-Haase, 2001; Tschannen-Moran et al., 1998). Thus, the GTE subscale may have a measurement issue which leads to low reliability and large variance in results.

Domain-specific measures based on locus of control. Regardless of the research showing conceptual issues (e.g., Guskey & Passaro, 1994; Hoy & Woolfolk, 1993), questions regarding the factor structure (e.g., Soodak, & Podell, 1996), and measurement validity and reliability concerns (e.g., Coladarci & Fink, 1995; Henson et al., 2001), the creation of domain-specific measures based on the TES began to emerge in the field of teacher self-efficacy. The compilation of teacher self-efficacy instruments that build upon previously flawed measures provides a disparity in the field; the move towards domain-specificity in measurement is one that aligns with Bandura’s original tenets of self-efficacy.

The most frequently used domain-specific measure based on the TES is the Science Teaching Efficacy Belief Instrument (STEBI) Forms A and B created by Riggs and Enochs (1990). Form A was created to measure the self-efficacy and outcome expectancy of in-service practicing elementary science teachers, whereas Form B was designed to assess these same beliefs in pre-service elementary science teachers.

Although the STEBI was created based on the flawed TES measure (Guskey & Passaro, 1994), many subsequent scales were developed based on the STEBI (i.e., Microcomputer Utilization in Teaching Efficacy Beliefs Instrument [MUTEBI], Enochs et al., 1993; Mathematics Teaching Efficacy Beliefs Instrument [MTEBI], Enochs et al., 2000; Table 2). Additionally, most of these self-efficacy measures focused in the context of elementary school (e.g., STEBI, Riggs & Enochs, 1990), domain of science (e.g., Self-Efficacy Beliefs about Equitable Science Teaching and Learning scale [SEBEST], Ritter, Boone, & Rubba, 2001), or were designed for the target population of pre-service teachers (e.g., Teaching Science as Inquiry [TSI], Smolleck, Zembal-Saul, & Yoker, 2006).

Secondary (context) specific measures based on locus of control. Two instruments based on the TES were found specifically designed to assess the perceived self-efficacy of high school teachers. The first (Barros et al, 2010) was designed solely for Physics teachers, and the second (Betoret, 2006) was part of a larger study but did not provide sufficient enough information to determine why specific self-efficacy questions were chosen, other than to state that they all loaded onto one factor (Table 1). Neither of these measures addressed the need for secondary specific measures in the field of literacy instruction, and both are aligned with locus of control rather than social cognitive theory.

Literacy (domain) specific measures based on locus of control. Only one literacy-specific measure was based off the STEBI, the Reading Teaching Efficacy Instrument (RTEI) by Szabo and Mokhtari (2004). This measure is task-specific for reading instruction; however it was designed for use with the target population of pre-service elementary teachers. Additionally, the items were created by changing the word "science" from the original STEBI to "reading," and so this scale has all of the fundamental issues of the STEBI itself. As the RTEI was designed for use with pre-service elementary teachers, the need to measure in-service secondary teachers' sense of efficacy in content area reading instruction still remains.

Summary of issues with measurement based on locus of control. As shown above there are two fundamental issues concerning teacher self-efficacy measures modeled after the TES. First, the TES was modeled after the RAND study which follows Rotter's (1966) locus of control theory and postulates people view experiences as internally or externally controlled. The difficulty in the TES framework is that Gibson and Dembo (1984) define self-efficacy and outcome expectancy as intertwined and their factors align more with the external versus internal locus (Guskey & Passaro, 1994) rather than what Bandura (1977, 1986) articulated as one's belief in his or her abilities to carry out a certain task versus the consequences such actions will produce. Therefore, the TES approaches teacher self-efficacy as a general construct and measures teachers' beliefs about student outcome expectations, instead of how Bandura originally conceptualized self-efficacy as a task and domain specific judgment about one's abilities to produce a certain outcome. As stated in a literature review by Klassen and colleagues (2011), almost one-third of the teachers' self-efficacy studies (*n* = 218) used some form of the TES.

Second, this measure is general and does not measure teachers' sense of efficacy in teaching in a particular context, domain, or task specific way; thus limiting its ability to causally predict any outcomes. Bandura (1997) articulated that self-efficacy is not a global construct, but rather a situational one and should be measured in that way. Furthermore, in order to link beliefs and outcomes it is important to measure beliefs within a specific context, domain, and task, in order to draw conclusions about outcomes based on those beliefs. Without specifying a specific teaching context, the TES simply measures a teacher's sense of efficacy in teaching in general, instead of at a specific task. And even though domain-specific instruments were created (e.g., STEBI, RTEI) to provide a more nuanced measure of teachers' efficacy, their development based off of the TES provided the same fundamental issues seen in the TES.

Furthermore, validity concerns have arisen concerning the two factor structure (Hoy & Woolfolk, 1993; Soodak & Podell, 1996) and the low internal reliability of the GTE factor (Henson, 2002; Henson et al., 2001). As Klassen and colleagues (2011) articulated, using flawed measures leads to misleading conclusions and diminishes the theoretical and predictive power of the field over time. Pajares (1996) echoed this sentiment stating that mis-measurement has "plagued" the field, and more accurate measurement of self-efficacy must occur in order for the field to progress. Although the domain-specific instruments that were created (e.g., STEBI, RTEI) advanced the field of self-efficacy, their development based off of the TES provided the same fundamental issues seen in the TES.

Social Cognitive Theory Framework

As various teacher self-efficacy instruments were created based off of Rotter's (1966) theory of locus of control, another strand of self-efficacy research was constructed grounded in Bandura's (1977) social cognitive theory. This theory postulates that people have human agency, can control what can happen to them, and thus can shape the outcome of their life (Bandura, 1977, 1986). Many teacher self-efficacy measures have been created based on Bandura's (1977) social cognitive theory. Bandura himself (1997) created a 30-item teacher self-efficacy scale with seven subscales: efficacy to influence decision making, efficacy to influence school resources, instructional efficacy, disciplinary efficacy, efficacy to enlist parental involvement, efficacy to enlist community involvement and efficacy to create a positive school climate (Tschannen-Moran et al., 1998). However, this measure was never published and outside of the discussion of it in Tschannen-Moran et al. (1998) no information regarding the factor structure, reliabilities, or validation was found in a review of the literature.

General measures based on social cognitive theory. One of the first teacher self-efficacy instruments created solely based on Bandura's (1977) conception of self-efficacy and the most used teacher self-efficacy measure is the Ohio State Teaching Efficacy Scale (OSTES) that was later named the Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Woolfolk Hoy, 2001). This scale is a general measure of teaching self-efficacy that was created and validated in three separate studies (i.e., Study 1, to explore factor structure; Study 2, to validate measure and explore the factor structure; Study 3, to confirm the factor structure and create a long and short version of the scale). Study 1 resulted in a factor matrix that would not converge and therefore items were revised and added to the measure. Study 2 yielded three factors: self-efficacy for student engagement, self-efficacy for instructional strategies, and self-efficacy for classroom management and offered validity (Table 2; Tschannen-Moran & Woolfolk Hoy, 2001). Study 3 confirmed the 3-factor structure and added the creation of a long (24 items) and short form (12 items) of the scale.

In study 2, Tschannen-Moran and Woolfolk Hoy (2001) used a mixed sample of pre and in-service teachers to validate the TSES and explore its factor structure. Convergent validity was established by assessing the correlations of scores on the RAND (r = .35 and .28, p < .01) and the TES (PTE, r = .48, p < .01; GTE, r = .30, p < .01). Discriminant validity was achieved by examining correlations on surveys of work alienation (r = -.31, p < .01) and pupil control ideology (r = -.25, p < .01; Tschannen-Moran & Woolfolk Hoy, 2001). Finally, the authors conducted factor analyses for both the pre-service and in-service teachers respectively and found that the three factor structure was appropriate among both groups. Thus, both versions of the scale were valid and the results of the data reliable.

Although validity was established across both the long and short versions of the measure, one main concern is the use of both pre and in-service teachers when validating the measure. In all three studies, the authors combined the results from pre-service and in-service teachers to determine which items to retain and to drop. This becomes problematic because Bandura (1986) hypothesized that mastery experiences impact one's self-efficacy above and beyond all other experiences, and pre and in service teachers have vastly different mastery experiences. Additionally, even though the authors confirmed the factor structure across each group independently in their final validation study, by splitting the sample, the sample sizes for each group becomes low (i.e., n = 111 pre-service teachers, n = 255 in-service teachers), which offers issues surrounding the number of items versus participants (Comfrey & Lee, 1992; Everitt, 1975). Thus, when examining the various steps of validation the authors engaged in, issues arise with the mixed sampling of pre-service and in-service teachers repeatedly being assessed in tandem.

Examining the TSES with pre and in-service teachers. As the TSES became more widely used in the field, studies emerged that attempted to revalidate the measure. Two studies (e.g., Duffin et al., 2012; Fives & Buehl, 2010) examined the TSES when used with pre and/or in-service teachers. Additionally, Fives and Buehl (2010) recounted past research that showed pre-service teachers demonstrated higher level of efficacy that decrease with experience. Do pre-service teachers actually incorrectly perceive their teaching self-efficacy simply because of a lack of mastery teaching experiences? Based on prior research, Duffin et al. (2012) suggested pre-service teachers may not have enough experiences to accurately respond to the task-specific self-efficacy items.

In their study, Fives and Buehl (2010) conducted a factor analysis with 372 participants (n = 102 in-service and n = 270 pre-service teachers). For practicing teachers, factor analysis on the long form and short form of the TSES yielded the same three factor solution found by Tschannen-Moran and Woolfolk Hoy (2001) with three exceptions where items factored higher on another factor than the original scale (Fives & Buehl, 2010). With pre-service teachers, the same analyses were conducted and factor analysis on the long and short form of the TSES revealed a single factor (Fives & Buehl, 2010). When the authors attempted to use the three factor solution, the results were not meaningful and the items did not load onto the correct factor. This suggests that a one factor solution may be more encompassing for pre-service teachers than the original three factor solution (Table 2).

Duffin and colleagues (2012) looked at two competing models of the TSES for use with pre-service teachers; study one (n = 272 pre-service teachers) used confirmatory factor analysis to extract one factor model and study two (n = 180 pre-service teachers) used confirmatory factor analysis to extract the original three factor model. Duffin and colleagues (2012) noted that such a high correlation among the three-factor solution (r = .86 and higher) demonstrated a one-factor solution was a better model fit. Expanding on this, the authors articulated that pre-service teachers were unable to distinguish between the different components of teaching (i.e., the three factors) captured by the TSES and thus all items loaded onto one factor (Duffin et al., 2012). Fives and Buehl (2010) and Duffin and colleagues (2012) both stated that researchers using the TSES with pre-service teachers should first use factor analysis to find the best factor structure before proceeding with analyzing their data.

Another study that focused on revalidating the TSES was conducted by Klassen and colleagues (2009) that used the short 12-item version of the instrument to measure teachers' sense of efficacy in 1212 teachers across five countries: Canada, Cyprus, Korea, Singapore, and the United States. The study provided five key insights into use of the TSES in varying settings. First, the three factor structure described by Tschannen-Moran and Woolfolk Hoy (2001) was more stable than a one factor structure tested by the researchers; second, internal consistency was reached across settings; third, the TSES showed measurement invariance across groups of the same culture by type of school setting (e.g., elementary versus middle/high school) thus teachers in different school settings had similar self-efficacy beliefs to others in the same school setting, but varying self-efficacy beliefs from those in a different setting; fourth, the TSES showed measurement invariance across cultural groups; and finally, the TSES supported the researchers hypothesized relationship between teaching self-efficacy and job satisfaction (Klassen et al., 2009). Although Klassen and colleagues (2009) reaffirmed the 3-factor structure of the TSES, all participants in the study were in-service teachers; a population that had accurately discriminated between the 3-factors in other studies (e.g., Fives & Buehl, 2010).

As researchers found differing results from their revalidation of the TSES (e.g., Duffin et al., 2012), the issue of validating instruments using the target population of interest becomes necessary. Moreover, when studies are conducted using a mixed sample of teachers with different mastery experiences, the factor structure of the instrument may not remain consistent across groups. Therefore, validation of instruments should never include a mixed sample of teachers with various mastery experiences as this can deter from the validation of the scale. More specifically, measures should be written for a specific population (e.g., secondary teachers, pre-service teachers, elementary teachers) with the contextual factors inherent to that population addressed.

Literacy-specific measures based on social cognitive theory. Many instruments based on Bandura's (1997) framework have been developed in an attempt to measure teachers' sense of efficacy through a social cognitive lens. As the field of teaching self-efficacy has progressed, so has the emergence of context, domain, and task-specific instruments. A search of the literature on teacher self-efficacy measures using Bandura's (1977) social cognitive theory identified four domain-specific literacy measures: two measures designed for teachers of English Language Learners (ELL) and two for pre and in-service reading/literacy teachers.

Teachers of ELLs. Two scales were created for teachers' of ELLs: the Exceptional Children who are English Learners (EXCEL) Teacher Inventory by Paneque and Barbetta (2006) and the English Language Teaching Efficacy Instrument (ELTEI) by Akbari and Tavassoli in 2014. Both of these scales are specific to literacy instruction, but each focuses on the teaching of the English language to ELLs and not to reading instruction to the general population. With the creation of the EXCEL, Paneque and Barbetta (2006) reported very few statistics in their results (Table 2) and did not provide validity beyond construct validity.

In their exploration of the factor structure of the ELTEI, the authors utilized a sample of 206 English language teachers and found a 7-factor structure; however fundamental issues arise when examining the construction of the instrument. First, the sample was very small and was utilized for both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA); a practice that poses methodological threats (Osborne & Costello, 2005). In addition, the authors provided no convergent or divergent validity to demonstrate their measure was indeed capturing what they intended. Thus, each of these scales has inherent validity concerns and reliability of scores must be assessed using replication studies.

Pre and In-Service Teacher Literacy Measures. Two measures were found that specifically examined teachers' sense of efficacy in literacy instruction: the Reading Teachers' Sense of Efficacy Scale (RTSES) by Haverback (2007, 2009) and the Teachers' Sense of Efficacy in Literacy Instruction (TSELI) by Tschannen-Moran and Johnson (2011). The RTSES was based on the TSES wherein the author added the term "reading" into the items to make them specific to reading instruction. It was designed specifically for pre-service teacher candidates sense of efficacy in literacy instruction based on tutoring and their field experiences (Haverback 2007; Haverback & Perault, 2011); and so it cannot be used with in-service teachers.

The second measure, the TSELI, was designed for in-service teachers; however it was validated using a primarily elementary sample and some items within the scale are more inherent in elementary contexts. In the validation study, 76% of the sample was elementary teachers with 24% being middle school teachers (Tschannen-Moran & Johnson, 2011). Additionally, since the sample is not broken down by teaching level and then subject area, it is difficult to know exactly what subjects the middle school teachers taught; especially as 14.7% (n = 85) of the sample reported teaching "other" subjects outside of the core content areas (i.e., science, social studies, math, English). Furthermore, some of the items on the scale only lend themselves to primary teachers (e.g., "To what extent can you implement word study strategies to teach spelling"); a practice not used in secondary classrooms. Thus, even though the instrument is literacy-specific for in-service teachers, it does not provide an appropriate framework to measure secondary teachers' sense of efficacy in teaching literacy instruction.

Need for New Measurement

As shown above, both frameworks (locus of control and social cognitive theory) struggle with mis-measurement, flawed replications of previous instruments, and, in some cases, poor validation methods. With the TES (Gibson & Dembo, 1984), there was a conceptually flawed alignment with Bandura's (1977) definition of self-efficacy and a general measurement of teacher self-efficacy. Modeled off of the TES, the STEBI, MUTEBI, MTEBI, and the RTEI exhibited similar conceptual and statistical issues found with the TES and all are built upon Rotter's locus of control theory.

In 2001 Tschannen-Moran and Woolfolk-Hoy created the TSES, the first published self-efficacy scale based on Bandura's (1986) social cognitive theory. Their instrument has no conceptual issues, but contradictory findings have occurred when using the instrument with pre and in-service teachers (Duffin et al., 2012; Fives & Buehl, 2010) and its general nature provides little predictive power. The literacy-specific instruments based off of social cognitive theory, EXCEL, ELTEI, and TSELI, all represent different domain specific instruments, but none measure secondary teachers' sense of efficacy in reading instruction. Moreover, the need to measure this population is still a fundamental concern for the field.

Measures must be created that are context (i.e., elementary or secondary), domain (i.e., content or subject area), and task (e.g., content-area literacy instruction) specific so that the target population (i.e., pre-service or in-service) is asked to assess their self-efficacy in regards to their mastery experiences in teaching within each of these areas. This is especially important in the area of content-area literacy instruction as the number of struggling readers performing below grade level on reading achievement tests has been steady over the past 20 years (Biancarosa & Snow, 2006; Nations Report Card, 2013). If secondary teachers are tasked with preparing adolescent students to be college and career ready, they must engage in reading instruction at the secondary level in an attempt to increase students' reading achievement. Therefore, new self-efficacy measures must be created to assess teachers' sense of efficacy for reading instruction in an attempt to link these beliefs to teacher practices and finally student reading achievement.

Additionally, new measures should be developed that align with Bandura's (1977) conception of self-efficacy. First, instruments must be developed specifically to assess the self-efficacy beliefs of secondary teachers' as their teaching context is vastly different than those of elementary teachers. Second, measures must be developed focusing on a specific domain or subject area because teachers' self-efficacy can change from domain to domain. Third, instruments must be designed to assess the particular task of content-area reading instruction. Fourth, when validating instruments and studying teacher self-efficacy it is essential to focus on one group (i.e., pre-service or in-service teachers) at a time since mastery experiences may influence teacher self-efficacy (Bandura, 1986, 1997), and these two groups of teachers have distinctly different mastery experience opportunities. Finally, methodological and psychometric procedures must be accurately addressed throughout construction, validation, and replication (for reliability) of instrumentation.

Context (secondary) specificity. The need for a secondary content-specific measure also remains. As shown in Klassen and colleagues (2010) review of the literature that found only 33 (15%) of the 213 studies examined were focused primarily on secondary teachers. The need to study secondary teachers exclusively is based on the idea that context (i.e., elementary or secondary) plays an integral role in teachers' self-efficacy. Specifically, secondary teachers teach in a different context than elementary teachers and it becomes detrimental to compare or equate the teaching self-efficacy of elementary teachers with that of secondary teachers. For example, elementary teachers are expected to teach one to two groups of students in one to two content areas, whereas secondary teachers teach three to five groups of students within one content area; therefore, when measuring the self-efficacy of teachers, it becomes necessary to design instruments that encompass the demands of their particular context.

In the current review of the literature there was no measure designed solely for use with secondary content area teachers in relation to literacy instruction. Furthermore, few studies (n = 9) included samples of primarily secondary teachers with no mixed sampling of elementary teachers. Yet most of these studies used general measures of teaching self-efficacy or researchers created a scale for use within their specific study without overt mention of item creation, validity, and reliability (Table 1). Thus, there is currently no domain-specific secondary-specific self-efficacy measure for researchers to utilize when studying an exclusive sample of high school teachers.

Domain (content areas) specificity. In accordance with Bandura's (1986, 1997) definition of self-efficacy, perceived self-efficacy beliefs are domain specific, and measurement should reflect this level of specificity. Although the TSES (Tschannen-Moran & Woolfolk Hoy, 2001) is based on Bandura's social cognitive theory framework of self-efficacy, the overall measure itself only accounts for teachers' general self-efficacy. Without domain or subject area specified, researchers cannot determine what subject teachers are referring when responding to self-efficacy items. Thus, if a teacher teaches multiple subject areas (domains) it is impossible for researchers to determine from which subject area teacher are responding when using a general self-efficacy measure. Therefore, it becomes essential for researchers to develop measures that align with Bandura's social cognitive framework of self-efficacy, in addition to the domain specificity required to predict outcomes from behaviors.

Task (content-area literacy instruction) specificity. Additionally, based on social cognitive theory, a teachers' sense of efficacy is dependent upon the task they are engaged, as some tasks are more difficult than other (Bandura, 1986). As the need for effective literacy instruction at the secondary level has become increasingly more important, due to the number of struggling readers in the US, understanding teachers' sense of efficacy beliefs regarding this type of instruction is necessary. Although literacy-specific self-efficacy measures have been created (i.e., TSELI, RSTES, EXCEL, and ELTEI) none of these measures address secondary content area teachers' self-efficacy in literacy instruction. The TSELI was created and validated using mostly (76%) elementary teachers, and has some items that are only specific to teaching at the elementary level (e.g., spelling; oral reading). The RSTES was written for use with pre-service teacher candidates, and was not meant for use with in-service teachers. Both the EXCEL and ELTEI are for use with teachers teaching the language of English to non-native speakers. Based on a review of the literature, the necessity for a content-area literacy-specific measure is still prevalent.

Target population and mastery experiences. A final area of concern in the current measurement of teacher self-efficacy is in the improper use of mixed samples of teachers (e.g., using pre and in-service teachers within the same study/for validity and reliability). As shown by research (e.g., Fives & Buehl, 2010) pre service and in-service teachers do not respond to self-efficacy measures in the same way; as Bandura’s (1986, 1997) postulated that mastery experiences influence one’s perceived self-efficacy above and beyond all other forms of persuasions. Also, because in-service teachers are actively engaged in teaching (mastery experience), they may view their self-efficacy differently than pre-service teachers sitting in an educational course with no real teaching experience.

Furthermore, studies have shown that it is incorrect to draw assumptions and connections between the self-efficacy of pre service and in-service teachers, as these self-efficacy judgments are created in two distinctly different spaces. Research has shown this trend, yet one of the most commonly used teacher self-efficacy scales used a sample of both pre and in-service teachers when creating, validating, and providing reliability for the measure and scoring (e.g., Tschannen-Moran & Woolfolk Hoy, 2001). Thus, when creating new measures it is important that studies focus specifically on one group (i.e., pre-service or in-service teachers) at a time to more accurately interpret factor structures, and provide a clearer picture of perceived teacher self-efficacy.

Literacy-specific teacher self-efficacy measure. With the insurgence of the CCSS and the push for literacy instruction in the content areas to address poor adolescent reaching achievement, secondary teachers are now faced with the task of infusing literacy instruction in their classrooms. As teachers, confront this new task, the opportunity arises for researchers to measure teachers' beliefs about their ability to effectively execute literacy instruction. Thus, the need arises for researchers to develop instruments to ascertain these context (secondary), domain (content area), and task (content-area literacy instruction) specific self-efficacy beliefs. By assessing secondary teachers' self-efficacy in literacy instruction, researchers can more effectively tie teaching efficacy to student outcomes, determine what makes effective and ineffective teachers, and address teachers' professional actions in relation to their perceived self-efficacy.

Based on the literature surrounding content-area literacy instruction in the secondary classroom, three specific areas were included in the development of the measure to assess teachers' sense of efficacy in a more nuanced fashion. Each of these areas has been shown through research to influence adolescent students reading proficiency in an attempt to decrease the number of struggling readers in the US. The three areas of content-area literacy instruction are: using reading strategy instruction (e.g., Vacca, 2002), motivating students to read (e.g., Guthrie & Wigfield, 2000), and teaching struggling readers (e.g., Biancarosa & Snow, 2006). Each of these three areas seeks to examine a specific aspect within the task of content-area literacy instruction at the secondary level and teachers' sense of efficacy in these three areas are important for secondary content-area literacy instruction.

Self-efficacy for using content-area reading strategy instruction. At the crux of content-area literacy instruction is the idea that teachers can instruct students on using a generic set of reading strategies to increase students' comprehension and overall ability to build knowledge from texts (Vacca, 1998; 2002). Examples of content area reading strategies include: activating background knowledge, questioning the text, drawing inferences, and creating mental images. Additionally, as Afflerbach and colleagues (2008) describe, reading strategies are deliberate attempts to control and modify a reader's ability to decode text, understand words, and construct meaning. The purpose of these strategies is to give students, especially those that are struggling a way to make meaning from difficult secondary texts; and offer above-grade readers additional ways to build comprehension. Furthermore, the importance of utilizing strategy instruction is demonstrated by Guthrie and colleagues (2000) who found that reading strategy instruction is correlated with students' intrinsic motivation for learning. Thus, strategy instruction not only leads to enhancements in student comprehension and subsequent achievement, but may also influence students' motivation to read as they offer students a means by which to actively engage with challenging texts, regardless of content area.

Content-area strategy instruction often follows the GRR model (Pearson & Gallagher, 1983); because when students experience explicit instruction on the definition and use of a specific skill, see teacher modeling, engage in guided practice, and conclude with independent practice, they are more likely to become proficient at the said skill (Biancarosa & Snow, 2006; Nokes & Dole, 2004). If a teacher does not believe that he/she is competent enough to incorporate reading strategy instruction in their classroom, thus having low self-efficacy, than research has shown that he/she will not engage in strategy instruction. Therefore, it is important for researchers to capture teachers' specific sense of efficacy in content area reading strategy instruction to more effectively design professional development opportunities. The professional experiences should reflect vicarious (e.g., a literacy coach modeling the GRR model of strategy instruction) and mastery (e.g., the teacher can practice using the GRR model of strategy instruction) experiences in order to increase teachers' self-efficacy beliefs about content area strategy instruction and consequently their professional practices in using content area reading strategy instruction.

Self-efficacy for motivating students to read. Students must be motivated to read in order to successfully engage in the reading process and foster comprehension. Explicit reading strategy instruction is one way to motivate even the most struggling readers to read within the content areas, as these strategies help to build comprehension and make meaning of secondary texts. In congruence, students must be motivated to read in order to engage with texts in a manner that promotes comprehension and understanding. Guthrie and Wigfield (1999) articulated how constructing meaning while reading is in itself a motivating act. Thus, it becomes important for teachers to help struggling readers construct meaning as they read, through the use of content area reading strategy instruction, in order to enhance student's motivation for reading.

In relation to instruction, teachers must feel efficacious in providing a motivating environment from which to engage students in reading in their content area, since student's motivation is strongly influenced by the experiences they have at school

(Wigfield, Guthrie, Tonks, & Perencevich, 2004). As students' primary experiences are with their teachers, external student motivation comes from the interactions they have with their teachers. In order to decrease the number of struggling readers in the US, secondary teachers must have high self-efficacy for using content-area reading strategies, motivating students to read, and teaching struggling readers in their content area; leading to an increase in their professional actions within these three areas.

Self-efficacy for teaching struggling readers. Coupled with the need for strategy instruction and motivating students to read at the secondary level, teaching struggling readers is of great importance to secondary instruction. As noted by Biancarosa and Snow (2006), 70% of high school students are considered struggling readers (i.e., they read at least two grade levels behind and have trouble with comprehension and making meaning from text). Most often struggling adolescent readers are marginalized readers whom are not connected to their classroom literacies and may have language or cultural practices different from those expected in schools (Moje, Young, Readence, & Moore 2000). Furthermore, struggling readers often become disenfranchised from education and schooling due to their inability to comprehend text and construct meaning from content materials.

Even though not all struggling readers are the same (e.g., some struggle with content, others struggle with comprehension), in order to effectively increase reading comprehension and achievement in the US, secondary teachers must have an increased sense of efficacy in helping these struggling readers. As Tschannen-Moran and Johnson (2011) noted, a teacher who did not expect to be successful in literacy instruction would likely put forth less effort in the preparation and delivery of instruction and give up more readily as the students struggled. Thus, self-efficacy plays a pivotal role in how much effort and persistence a teacher shows when faced with difficult circumstances and lower performing students.

Summary

Much of the current teacher self-efficacy research has demonstrated that mis-measurement has continually plagued the field, from item development, lack of theoretical conceptualization to poor validity and reliability coefficients (Henson, 2002; Klassen et al., 2011). Therefore, it becomes important for context (i.e., primary or secondary), domain (i.e., content or subject area), and task-specific (e.g., reading instruction) self-efficacy measures to be created based on Bandura's (1977, 1986) social cognitive framework. Based on the adolescent literacy and teacher self-efficacy literature, an area that has been underexplored in research is secondary teachers' sense of efficacy in literacy instruction across the content areas.

As adolescent literacy achievement is a concern in the US with 64% of 8^{th} grade students reading at basic or below basic levels (Nations Report Card, 2013), the need to measure teachers' sense of efficacy for reading instruction becomes necessary. By accurately assessing these beliefs, researchers can determine what fosters or hinders teachers' use of reading instruction in an effort to increase successful literacy instruction at the secondary level. Thus, a new measure must be created that addresses secondary content-area literacy instruction: using content-area reading strategy instruction, motivating students to read, and teaching struggling readers. Additionally, this measure must be validated using only the target population to provide consistent internal reliability before use within the field.

Chapter Three: Methodology

The purpose of this study was to develop and validate an instrument to measure Secondary Teachers' Sense of Efficacy in Reading Instruction (STERI) across the content areas (i.e., science, social studies, mathematics, English). Although the instrument was meant to be used with all secondary teachers (i.e., grades 6-12) for the purposes of initial survey development and exploratory analyses, only high school teachers (i.e., grades 9-12) were included in this study to make the sample as homogenous as possible. Having a homoegenous sample helps to alleviate any extraneous variables that could potentially confound validation studies.

As survey development is a multi-step process, the initial development (e.g., literature review, writing of items) of the STERI measure occurred before Exploratory Factor Analysis (EFA). Therefore, this chapter first describes the methods by which items were created. Then, it addresses how I revised the items in the instrument using qualitative methods (i.e., expert review and cognitive interviews). Finally, it explains how I explored the structure, established internal consistency reliability, examined descriptive statistics, hypothesized relationships between STERI subscales, and established validity of the instrument using quantitative methods (e.g., EFA, Analysis of Variance [ANOVA], correlations).

Instrument Development

This study was structured around the six step approach described by Gehlbach and Brinkworth (2011), along with methodological suggestions for scale development (e.g., Crocker & Algina, 1986; DeVellis, 2003) and previous self-efficacy scale development research (e.g., Dellinger et al., 2008; Klassen, Yerdelen, & Durksen, 2013; Tschannen-Moran & Johnson, 2011). In Gehlbach and Brinkworth's (2011) model, they delineated the use of: a) a literature reviews; b) interviews with target population; c) a synthesis of the information from the literature review and interviews; d) item development; e) expert reviews; and f) cognitive interviews, all of which are meant to frontload validity before pilot testing occurs. Validity is an important component of instrument development as it refers to the degree to which an instrument accurately reflects the construct it is intending to measure (DeVellis, 2003). Based on the literature, I created a ten-step process model (Figure 1) to engage in multiple qualitative and quantitative methods in an attempt to provide an instrument that is a clear representation of the constructs of adolescent literacy instruction and teacher self-efficacy.

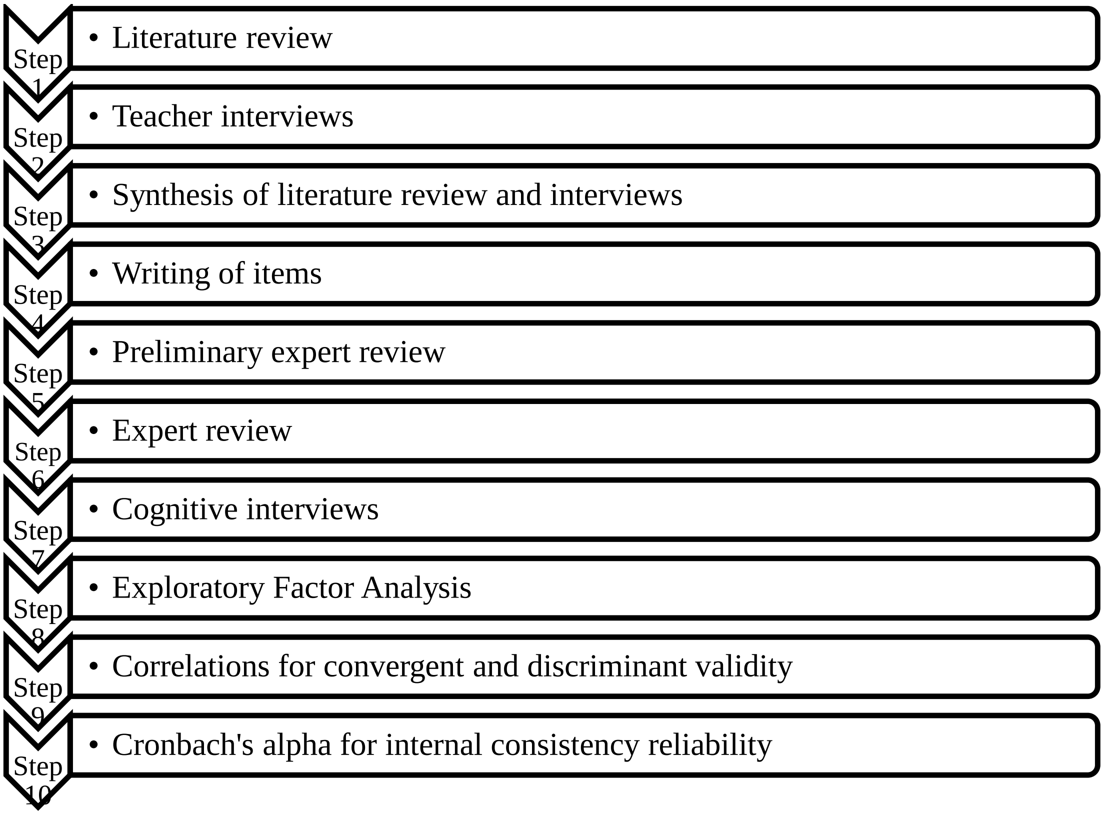

Figure 1. Ten step model based off of the work of Gehlbach and Brinkworth (2011).

Item Generation

One of the most important facets of survey development is the writing of items that accurately measure variations of the desired construct (DeVellis, 2003; Gehlbach & Brinkworth, 2011; Switzer, Wisniewski, Belle, Dew, & Schultz, 1999; Wheatley, 2005). Research has shown that many of the issues surrounding past self-efficacy measures come from poorly written items, items that cross-load on factors, and items that are politically charged leading to ceiling and basement effects (Tschannen-Moran & Woolfolk Hoy, 2001). To combat these validity issues, and following the survey development methodology described by Crocker and Algina (1986) and DeVellis (2003), I began by conducting a thorough review to find literature in which researchers: described their processes of teacher self-efficacy measure development, discussed issues with current teacher self-efficacy instruments, and/or focused on measuring teacher self-efficacy in relation to literacy. This approach is further supported by Soodak and Podell (1996) who argued that the qualitative soundness of item writing is essential to create measures that address the constructs they are intended to measure.

Additionally, many of the previously developed teacher self-efficacy scales began with a thorough review of the literature to form a pool of items to address the construct (e.g., Klassen et al., 2013; Paneque & Barbetta, 2006; Tschannen-Moran & Johnson, 2011). After completing a comprehensive literature review, I interviewed teachers from the target population to provide support for these findings. Gehlbach and Brinkworth (2011) discussed the use of interviews to support literature review findings and provide further evidence for the inclusion of factors. Further, I synthesized the findings from both the literature review and teacher interviews to formulate a comprehensive pool of items.

Literature review. In accordance with Gehlbach and Brinkworth's (2011) tenets, I looked at the literature with two specific goals: 1) to synthesize what content to include in item writing based on the adolescent literacy literature, and 2) to examine how other instruments had captured the construct of self-efficacy. During the literature review process, I synthesized the adolescent literacy literature and determined four key areas that needed to be included in the STERI. These areas were chosen because they represented key constructs within the content-area and adolescent literacy literature. The four areas were: self-efficacy for content-area reading strategy instruction (e.g., Vacca, 2002), self-efficacy for Motivating Students to Read (e.g., Guthrie et al., 1996), self-efficacy for Teaching Struggling Readers (e.g., Biancarosa & Snow, 2006), and self-efficacy for Using assessments (e.g., Moje et al., 2000). After deciding upon the inclusion of these four factors, major concepts from each area were recorded in Microsoft Word to help with item writing.

In addition, I created two tables organized by guiding theory (i.e., locus of control or social cognitive theory) in Microsoft Excel to examine the psychometric properties of pre-existing teacher self-efficacy measures. In the first table, I recorded the following from each scale: number of items, number and label of anchor scores, factor structure (i.e., number and name of factors), factor loadings, reported reliability, and variance explained (Table 2). In the second table, I recorded: sample size used, type of validity established (i.e., face, content, convergent, discriminant), and what scales were used to establish validity (Table 3). These tables were used to examine the teacher self-efficacy measures that have been created to date, and to help determine sample size requirements, number of items, and the process needed to accurately validate the STERI measure.

Table 3

Sample Sizes, Validity, and Instrumentation for Teacher Self-Efficacy Instruments

Instrument	**Sample Size**	**Face Validity**	**Content Validity**	**Convergent Validity**	**Discriminant Validity**
Measures based on Locus of Control Theory					
Teacher Locus of Control (TLC; Rose & Medway, 1981)	134 IS teachers	183 IS teachers		Locus of Control; Implementation of Educational Practices using observations	Crowne-Marlowe Social Desirability scale
Teacher Efficacy Scale (TES; Gibson & Dembo, 1984)	Phase 1: 203 IS teachers Phase 2: 55 PS teachers	90 IS teachers		Open-Ended measure of teacher efficacy	Teacher Aptitude Battery: Finding Useful Parts Test; Planning Test; Verbal Facility Test; Controlled Associations Test
Science Teacher Efficacy Belief Instrument (STEBI; Riggs & Enochs, 1990)	212 PS teachers		5 science educators	Subject Preference Inventory for 28 PS teachers	

Microcomputer Utilization in Teaching Efficacy Beliefs Instrument (MUTEBI; Enochs et al., 1993)	232 IS teachers		3 science and 2 microcomputer educators	2 researcher-derived questions in relation to experience and use
Mathematics Teaching Efficacy Beliefs Instrument (MTE; Enochs, Smith, & Huinker, 2000)	324 PS teachers			
Self-Efficacy Beliefs about Equitable Science Teaching and Learning (SEBEST; Ritter et al., 2001)	217 PS teachers	10 graduate students in science education	8 faculty members in science education, multicultural education, and self-efficacy	
Teaching Science as Inquiry (TSI; Smolleck et al., 2006)	184 PS teachers	190 PS teachers	6 versions of content validity: 6 faculty members and 3 graduate students; 9 experts; 3 faculty members; 4 faculty members; faculty members; faculty members	

Reading Teaching Efficacy Instrument (RTEI; Szabo & Mokhtari, 2004)	419 PS teachers	Group of teacher candidates	4 literacy faculty members, 3 reading specialists, 5 doctoral students in literacy education	Teacher Efficacy Beliefs Instrument (TEBI)	
Measures based on Social Cognitive Theory					
Ohio State Teaching Efficacy Scale (OSTES)/ Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Woolfolk Hoy, 2001)	Study 1: 146 PS teachers 78 IS teachers	Sample used for face validity			
	Study 2: 70 PS teachers 147 IS teachers			RAND; 10-item TES; Pupil Control Ideology	Work alienation survey
	Study 3: 103 PS teachers 255 IS teachers			RAND; 10-item TE	
Teachers' Efficacy Beliefs System-Self Form (TEBS-Self; Dellinger et al., 2008)	1437 IS teachers (Olivier, 2000)		45 professional educators		
	555 IS teachers (Bobbett, 2001)		45 professional educators		
	381 IS teachers (Dellinger, 2001)		45 professional educators		

Exceptional Children who are English Learners (EXCEL) Teacher Inventory (Paneque & Barbetta, 2006)	202 IS teachers	Review panel of 3 experts; 20 special education teachers	Review of the literature	
English Language Teacher Efficacy Instrument (ELTEI; Akbari & Tavassoli, 2014)	206 IS teachers	Interviews of 18 teachers	4 researchers/teachers who had previous experience in working on teacher efficacy	
Teachers' Sense of Efficacy for Literacy Instruction (TSELI; Tschannen-Moran & Johnson, 2011)	648 IS teachers	Field tested with 11 graduate student in literacy instruction	4 experts in the field of reading and literacy instruction	12-item TSES
Readings Teacher Sense of Efficacy Scale (RTSES; Haverback & Parault, 2011)	102 PS teachers	Random groups of preservice teachers over three semesters	7 educational specialists	

Note. IS = In-Service. PS = Pre-Service. Empty cells are data not reported in research article.

Teacher interviews. Upon completion of the literature review process, I conducted semi-structured interviews with the target population to help determine how they viewed adolescent literacy instruction at the secondary level. By engaging in interviews, researchers can determine if and how participants think about the focal construct, and then ask probing questions to establish if the participants' ideas match that of the literature (Gehlbach & Brinkworth, 2011). In regards to this study, participants were asked about their: thoughts about content area literacy instruction, sense of efficacy for using content area literacy instruction, and preparation for using content area literacy instruction.

Participants were two seventh grade social studies teachers from a diverse school district on the east coast that were part of a larger literacy intervention. Both participants were male, had at least three years of teaching experience in social studies, and held at least a bachelor's degree. The teachers' only experience with teaching literacy was through participation in the larger study. Semi-structured interviews were conducted by two graduate research assistants (i.e., one graduate student per teacher). These interviews were intended to capture teachers' beliefs after the implementation of the larger literacy intervention; however, I received permission from the principal investigators to include questions specifically targeted to aid in item writing for the STERI (Appendix B). Interviews lasted between 30-60 minutes and were audio recorded for analytical purposes.

Following Gehlbach and Brinkworth's (2011) framework, teachers were asked general questions about adolescent literacy instruction (e.g., "What knowledge do you think teachers need in order to incorporate reading comprehension in social studies?”). Then, based on their response, follow-up questions probed participants to expand on their initial responses (e.g., “You mentioned reading strategies; do you think having a strategy to link onto helps?”). Questions were also added to the protocol that asked teachers about their teaching self-efficacy in literacy instruction (e.g., “How confident or comfortable do you feel in teaching different strategies?”). As these interviews were part of a larger intervention study, for the purposes of this study, I only focused on the questions and responses that pertained to literacy instruction in social studies. A full list of questions added to the semi-structured interview post-intervention protocol can be found in Appendix B. Each teacher interview was transcribed and I analyzed the data to determine if the teacher’s responses mirrored what was found in the literature.

First, I recorded a comprehensive list of topics each teacher discussed in detail (i.e., those mentioned at least twice during the interview; Table 4). As these interviews were centered within a larger intervention study, some of the topics teachers discussed were more relevant to the examination of the intervention than for the purposes of confirming the adolescent literacy findings (e.g., issues with time, inter-disciplinary teaching).

Table 4

Topics that Arose from Teacher Interviews

Teacher 1	Teacher 2
Students reading below grade level	Students reading below grade level
Motivating students in social studies	Motivating students in social studies
Using strategies to aid in comprehension	Using strategies to aid in comprehension
Vocabulary	*Lack of teacher preparation in integrating literacy practices into social studies*
Projects	Using the textbook
Bringing in outside sources	Modeling
Working with ELLs and their struggles	Students in the classroom affect instruction
Modeling	What good readers do while reading
Issues with time	Reading as a big component in social studies
Using the textbook	Working with ELLs
Students in the classroom affect instruction	*Inter-disciplinary teaching*
What good readers do while reading	*Special vocabulary*
Reading as a big component in social studies	Student participation
Student participation	

Note. Italicized words are those topics mentioned by teachers more than once, but were not included in the synthesis of the literature on adolescent literacy instruction.

Next, to reinforce what was found in the literature I examined the topics mentioned to look for the concepts of: using content-area reading strategy instruction, motivating students to read, teaching struggling readers, and using assessments. From these interviews, I found support for three of the concepts identified in the adolescent literacy literature. For example, both teachers discussed: using strategies to aid in comprehension, modeling, reading as a big component in social studies, and explaining what good readers do while reading. All of these ideas matched the adolescent literacy literature regarding content-area reading strategy instruction. In addition, both teachers discussed: motivating students to read in social studies, student participation, and other motivating activities (e.g., bringing in outside sources, projects, using the textbook). These topics supported what the literature defined as motivational practices.

Finally, in the interviews both teachers discussed: ELLs and their struggles, students reading below grade level, and students in the classroom affecting instruction. All of these topics matched what the adolescent literacy literature defined as struggling readers. The fourth concept, using assessments, was only found in the literature but was included in the initial writing of items. From the aligning of teacher interview data to the adolescent literacy research, the four factors of the STERI were named: self-efficacy for Using Content-Area Reading Strategy Instruction, self-efficacy for Motivating Students to Read, self-efficacy for Teaching Struggling Readers, and self-efficacy for Using Assessments.

In addition, to provide further support for these four factors, I examined each transcript and found quotes that represented and supported the inclusion of three of the four areas (Appendix C, Table 5). Thus, I was able to combine concepts that the teachers discussed with what I had already found in the literature surrounding three of the four factors. In addition to these three factors, I initially believed assessment to be an integral part of adolescent literacy instruction based on the literature. Therefore, a fourth factor of assessment was incorporated into the STERI even though the teacher interviews did not address this factor.

Synthesis of literature review and teacher interviews. In combining the information from the literature review and teacher interviews, I noted that both teachers discussed: reading strategy use, student motivation and engagement, and struggling readers in detail during their interviews, a concurrent finding in the adolescent literacy literature. Therefore, I concluded that these three factors were an accurate representation of content area literacy instruction in the secondary classroom. A fourth factor of assessment was included in the initial STERI as it was believed to be important to content-area literacy instruction based upon a review of the literature. However, this finding was not supported by the teacher interviews. Then, I consulted Bandura's (2006) *Guide for Constructing Self-Efficacy Scales* to determine anchor scores and the proper procedure for writing items.

Based on Bandura's (2006) suggestions, all items were written using *can do* instead of *will do* to capture participants' perceived self-efficacy of their capabilities not their intentions. By writing items with this approach, participants are forced to examine the construct of self-efficacy as something that is occurring at this point in time, instead of somewhere in the distant future. Also, based on previous measurement issues and in accordance with Bandura's (2006) guide, items were written to represent tasks of various degrees of difficulty to reduce the chance of ceiling effects. Ceiling effects occur when all participants respond as highly efficacious because the items in the measure do not represent complex or difficult tasks. To combat ceiling effects, items in the STERI represent tasks that are not easily performed (e.g., "To what extent can you use assessments to help struggling readers in your content area?"), which may offer greater variability in scores.

Additionally, a 6-point anchor scale was utilized with phrases ranging from *not at all* to *completely*. This decision was based primarily on Bandura's (2006) assertion that scales using only a few steps are less reliable and often lose the differentiation needed to examine differences in responses. Furthermore, scales with too few anchor scores can lose their sensitivity, as participants are forced to choose between too few options and therefore their differences are not recorded (Bandura, 2006). Conversely, a scale with too many anchor scores can be cognitively taxing for participants (Miller, 1956). Although Bandura (2006) suggests a 0-100 scale in increments of 10 (i.e., an 11-point scale), a 6-point scale was chosen to offer participants enough choices without overwhelming them with response categories (DeVellis, 2012; Gehlbach & Brinkworth, 2011; Krosnick, 1999; Miller, 1956; Smith, Wakely, De Kruif, & Swartz, 2003).

Moreover, a 6-point scale was used to avoid a true middle anchor (i.e., a neutral response option) to force participants to make a decision, instead of gravitating towards the most neutral response category. Thus, a 6-point scale offers the range Bandura (2006) suggests and creates the condition where participants cannot simply mark a neutral response, but rather must articulate a higher or lower sense of efficacy for each item. Also, as stated in the literature, agree/disagree were not used as anchor scores as these are too cognitive taxing (e.g., Krosnick, 1999). Further, verbal labels were used in lieu of numbers, as numbers can have an implicit meaning for participants (e.g., Tourangeau, Rips, & Rasinski, 2000). Using verbal instead of numerical labels decreases the possibility of respondents lumping together numerical categories (i.e., seeing 1-3 as all representing "not at all" when only one was meant to reflect this response) and provides a clear distinction in each anchor score. The 6-point Likert scale with verbal labels consisted of: *Not at all; Very little; Somewhat; Moderately; Quite a bit; and Completely* and represented a continuum from 1-6 using verbal instead of numerical labels.

Writing of items. Based on the survey development literature, items were written to avoid: bias, ambiguous language, and multi-faceted question stems (e.g., DeVellis, 2003). Also, items were written to encompass each of the factors in their entirety (i.e., Using Content-Area Reading Strategy Instruction, Motivating Student to Read, Teaching Struggling Readers, and Using Assessments); therefore redundancy was prevalent among the initial pool of items. This process resulted in an exhaustive list of 121 items based around the four initial factors. Following Gehlbach and Brinkworth's (2011) framework, items were written to apply to every respondent by including "in your classroom/content area" to ensure that teachers could see themselves in each item. Further, these items reflected Bandura's social cognitive framework and related teachers' sense of efficacy to their personal, behavioral and environmental factors (e.g., To what extent can you motivate students to read in your content area?); instead of Rotter's external versus internal locus of control framework (e.g., How much of a student's motivation to read is outside of the teacher's control?).

Item writing began in January 2014 and the initial instrument included 121 items: 31 items for the Using Content-Area Reading Strategy Instruction factor, 30 items for the

Motivating Students to Read factor, 29 items for the Teaching Struggling Readers factor, and 29 items for a fourth factor of Using Assessments. From January to May, 2014 I met on a consistent (i.e., weekly or bi-weekly) basis with an educational psychology professor familiar with self-efficacy research and scale development to review and revise the items. During these meetings anticipated factors were defined, previous self-efficacy measures were discussed, items were extensively revised, and multiple items were dropped.

Additionally, the fourth factor of Using assessments was dropped during the revising process (March 2014) due to difficulties surrounding qualitative writing of items and the overall fit of the factor to the scale. More specifically, the fourth factor of Using Assessments became too heavily populated with items and created a distraction from the rest of the measure. Therefore, instead of having a factor labeled assessment, specific items were written that addressed assessment in relation to the other three factors. For example an item was added to the Teaching Struggling Readers factor that stated, "How confident are you in your ability to use a variety of assessments to help struggling readers in your content area?" Thus, the 29 items on the Using Assessments factor were deleted from the measure, and items related to assessment were added to the three remaining factors.

Over time, the total number of items reduced from 121 to 46. These deletions occurred because: 29 items were from the Using Assessments factor, 7 items were redundant, 14 items addressed constructs outside of self-efficacy, and 23 items addressed teacher preparation. After revisions, the total number of items was 46: 17 items for the Using Content-Area Reading Strategy Instruction factor, 10 items for the Motivating

Students to Read factor, and 19 items for the Teaching Struggling Readers factor (Appendix D).

Item Revision and Reduction

In order to examine how items were interpreted, and to reduce the number of items (n = 46) even further, experts in the fields of adolescent/content-area literacy and self-efficacy (n = 10) were asked to review the initial STERI instrument. This served two purposes: to help establish face and content validity and to reduce the number of items on the scale before exploratory testing. After expert review, two rounds of cognitive interviews (n = 12) occurred to further establish face validity and revise items before large-scale quantitative analyses.

Expert review. As recommended in scale development (e.g., Crocker & Algina, 1986; DeVellis, 2003; Gehlbach & Brinkworth, 2011), experts can be used to establish content validity during measure development. Content validity is the extent to which the items adequately represent the specific construct, and can be established when experts in the field can agree the items in the scale represent the target construct (Crocker & Algina, 1986; DeVellis, 2003). For the purposes of this study, participants were considered experts if they had research experience in the adolescent/content-area literacy or self-efficacy fields. Experts from both fields were used to establish content validity as the STERI represents self-efficacy measurement as well as the three hypothesized factors related to adolescent literacy (i.e., Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers).

Preliminary expert review of items. In order to pilot the STERI for expert review, six doctoral literacy students from a large public university were recruited to initially review and reduce the number of items. In addition to being doctoral literacy students, the participants were all current or former teachers and therefore could provide face validity for the instrument. Face validity is a qualitatively subjective assessment of whether or not the instrument "looks like" it is going to measure what it is intended to measure (Rubin & Babbie, 2007). This assessment is done with a sample of participants from the target population whom examine the instrument to determine if they believe it will measure what it is intended. The purpose of using these participants as preliminary experts was to further revise, and reduce the number of items before sending the measure to a second round of experts in the fields of adolescent/content-area literacy and self-efficacy measurement.

The STERI was redesigned specifically for expert review by creating two columns within the measure, one for relevance and one for clarity (Figure 2) that participants could use to rate the items. In order to ensure the participants were clear on the constructs being measured, definitions were provided for each of the three factors (e.g., "Self-efficacy for Teaching Struggling Readers, definition: How confident content area teachers feel when teaching struggling readers. Struggling readers are those students who read at below-grade reading levels and/or have difficulties connecting to literacy texts/processes"). In addition, a 4-point rating scale was used with responses ranging from "1=Not Acceptable (major modifications needed)" to "4=Exceeds Expectations (no modifications needed)." Thus, participants rated each item on the 1-4 scale for both its relevance and clarity. Therefore, two scores were reported for each item (i.e., one for relevance and one for clarity). Participants were also given the opportunity to provide qualitative notes on any items, factors, or the instrument as a whole. Qualitative notes were given in three forms: within individual items through tracked changes in Microsoft word, as comments throughout the measure, and/or at the end of each item in the space provided.

<table>
<tr><td>This survey is designed to help capture teachers' beliefs about literacy instruction within the various content areas (Science, Math, History, or English). Please read the following items and rate each on clarity of wording and relevance to construct. Feel free to suggest revisions, deletions, or additions of items.</td><td colspan="3">Score
1=Not Acceptable (major modifications needed)
2=Below Expectations (some modifications needed)
3=Meets Expectations (no modifications needed but could be improved with minor changes)
4=Exceeds Expectations (no modifications needed)</td></tr>
<tr><td>For color-coded items, please choose best item.</td><td>Clarity</td><td>Relevance</td><td>Suggested Revisions</td></tr>
<tr><td colspan="4">Self- Efficacy for Helping Struggling Readers
Def: How confident content area teachers feel when teaching struggling readers. Struggling readers are those students who read at below-grade reading levels and/or have difficulties connecting to literacy texts/processes.</td></tr>
<tr><td>To what extent can you diagnose why a student is struggling with reading in your content area?</td><td></td><td></td><td></td></tr>
<tr><td>To what extent can you model good reading in your content area to help struggling readers?</td><td></td><td></td><td></td></tr>
</table>

Figure 2. STERI redesigned for expert review.

The STERI was sent to the six doctoral literacy students as a Microsoft Word document via email. Additionally, the measure for the doctoral literacy students contained three pairs of items that were extremely similar in wording. These items were color coded (e.g., each pair received a different highlighter color) and respondents were instructed to choose the best item of the pair based on clarity and relevance. This process was done to delete three of the six items that were extremely similar. Participants were given one month to review the items and rate them on the two aspects: relevance and clarity.

Upon receiving the reviews from participants, all quantitative data were entered into SPSS and all qualitative responses recorded in Microsoft Word. Then mean ratings and standard deviations for each item based on a) relevance and b) clarity were analyzed. Next, the qualitative responses were examined and revisions to item wording were undertaken. Finally, the quantitative and qualitative ratings were examined across all participants and final item revisions and reductions occurred to prepare the measure for further expert review.

From this initial round of expert review, the total number of items on the STERI was reduced from 46 to 31. The decision to remove items was based on at least half (n = 3) of the experts agreeing that an item was redundant or low on clarity/importance. Of these deletions, 10 items were removed based on repetitiveness, and five items were deleted due to confusing wording, low clarity (mean rating of less than 3.0), and/or low relevance (mean rating of less than 3.0) to the construct (Appendix E). All remaining items in the STERI had a mean relevance rating of 3.0 or higher and a mean clarity rating of 3.0 or higher. Therefore, all items remaining in the STERI met or exceeded expectations for both relevance and clarity.

Additionally, the number of items per factor were: self-efficacy for Using Content-Area Reading Strategy Instruction, eight items; self-efficacy for Motivating Students to Read, nine items; and self-efficacy for Teaching Struggling Readers, 14 items. Further, based upon suggestions by the majority of experts (n = 4) items that included the phrase "good readers" were changed to "effective readers." Finally, the stem of items was changed from "How much can you/To what extent can you" into "How confident are you in your ability to" based on the majority of experts (n = 5) articulating that the stem of the items should be consistent for readability. These revisions were made to help with the clarity and meaning of items.

Second expert review of items. In order to establish content validity, experts in the fields of content-area literacy and self-efficacy were asked to review the measure to help provide further validity. Content validity involves providing evidence that the content is relevant, representative, and all-encompassing of the construct of interest (DeVellis, 2003; Gehlbach & Brinkworth, 2011). Content validity can be a qualitative rating of items on clarity and importance to the construct being measured. Two experts from content-area literacy field, Dr. Nancy Frey and Dr. Maryann Mraz, and two experts from the self-efficacy field, Dr. Anita Woolfolk Hoy and Dr. Ellen Usher, reviewed the STERI.

Dr. Nancy Frey is a Professor of Literacy in the School of Teacher Education at San Diego State University. Dr. Frey's research interests include reading and literacy, assessment, intervention, and curriculum design. Dr. Maryann Mraz is a Professor in the Reading and Elementary Education Department at the University of North Carolina at Charlotte. Dr. Mraz's professional interests include literacy coaching, professional development, early literacy, and content area literacy. Dr. Anita Woolfolk Hoy is a Professor Emeritus in the Educational Psychology and Philosophy Program within the

College of Education and Human Ecology at The Ohio State University. Dr. Woolfolk Hoy has published research in the areas of teacher cognition, student perceptions of teachers, teachers' beliefs, student motivation, and the application of educational psychology to teaching. Dr. Ellen Usher is an Associate Professor and Director of the P20 Motivation and Learning Lab at the University of Kentucky. Dr. Usher's research focuses on the sources and effects of beliefs of personal efficacy from the perspective of social cognitive theory.

These experts were chosen based on their work within their respective fields, but also because they represented different perspectives within their fields (i.e., as Fisher and Frey conduct research together, to get a different view Dr. Mraz was the second reviewer; Woolfolk Hoy and Tschannen-Moran conduct research together, so Dr. Usher was chosen to be the second reviewer). Thus, I was able to get a broader perspective of the field using researchers whom did not often publish together. These experts were recruited in-person (i.e., at the 2014 Advances in Educational Psychology Conference) or via email and all gave permission to use their names within this study (Appendices W-Z). After experts agreed to review the STERI, the measure was sent to them via email. In addition, each participant was told to review the STERI based on their expertise in their respective field (i.e., self-efficacy or content-area literacy). These experts used the same STERI form that was used for the preliminary experts and rated each item on relevance and clarity, and provided qualitative feedback.

Analysis followed the same procedure completed with the doctoral student review. After receiving the reviews from all four expert reviewers, all quantitative data were entered into SPSS and all qualitative responses recorded in Microsoft Word. As three out of the four reviewers gave qualitative notes item by item (i.e., specific comments were made next to items not in a summary paragraph) these notes were all condensed into one document. Then mean ratings for each item based on a) relevance and b) clarity were analyzed and added into the qualitative document. Next, the qualitative responses were examined and revisions to item wording were undertaken. These data were used to determine which items needed to be revised or dropped from the measure before engaging in cognitive interviewing. The results of these analyses are discussed in detail in Chapter 4.

Cognitive interviews. Based on recommendations for survey development, cognitive interviews were used to determine how respondents from the target population (i.e., high school content areas teachers) interpreted items within the STERI (Campanelli, 1997; DeMaio & Rothgeb, 1996; Gehlbach & Brinkworth, 2011; Willis, 2005). According to Willis (2005), cognitive interviewing can aid researchers in how the target audience will understand, process, and respond to survey items; with a key emphasis on the potential breakdown of this process. Using cognitive interviews aided in formulating a better understanding of the process by which teachers were reading, engaging, and responding to the items. This also offered further face validity for the instrument as these teachers represented a subset of the target population.

As shown in Chapter 2, developing qualitatively sound items that function as the researcher intended has been an issue for some of the previous self-efficacy measures (e.g., TES; Gibson & Dembo, 1984). For survey researchers, the development and subsequent evaluation of items is of critical concern. Cognitive interviewing has emerged as a method to help identify and correct any problems with survey items (Beatty & Willis, 2007; Campanelli, 1997; Willis, 2005). Although there is no one standardized method by which to engage in cognitive interviewing, all common applications for cognitive interviewing entail participants be engaged with draft survey items and provide verbal information about the items. This information is then used to evaluate the quality of the items and determine whether the items are generating the information as the author had intended (Beatty & Willis, 2007).

Two main methods of cognitive interviewing have emerged over time: think-aloud and verbal probing. Think-aloud procedures call for the interviewer to read the question aloud and the respondent thinks through their decisions aloud. Verbal probing is used by asking participants targeted questions after they respond to an item to explore their beliefs and understandings in an in-depth manner. According to Willis (2004): the advantages to think-alouds are: a) freedom from interview-imposed bias; b) minimal interview training requirements; and c) open-ended format; whereas, the disadvantages are: a) subjects must be trained in the process; b) subject resistance; c) cognitive demands on subjects; d) subjects straying from the task; and e) bias in subject information processing. In contrast, the advantages to probing are: a) control of the interview; and b) ease of training the subject; and the disadvantages are: a) probes can be artificial in nature; and b) bias from the researcher can influence responses (Willis, 2005). Based on the number of disadvantages and the unnatural aspect of asking participants to think aloud while answering survey items, verbal probing was used as the primary method to engage in cognitive interviewing.

Following the seminal work of Belson (1981), participants were asked specific questions to explore their: interpretations of questions, responses to items, and identifications of any reporting errors. Reporting errors occur when participants do not interpret the items as they were originally intended which can lead to items cross-loading or not loading onto the intended factor. Therefore, cognitive interviewing has become an analytical tool by which survey designers can gain qualitative insights into how items are being interpreted by the target population.

The process of concurrent verbal probing, a type of verbal probing, consists of: the interviewer asking the survey question, the participant answering the survey question, the interviewer probing the participant, and the participant answering the probing question (Willis, 2005). Concurrent verbal probing is a process where the interviewer seeks to gain detailed responses regarding survey items by engaging the participant in a consistent cyclical process throughout the interview (Figure 3). In this study, participants were asked probing questions about their comprehension, confidence ratings, and ability to paraphrase questions (DeMaio & Rothgeb, 1996). For example, based on the work of Karabenick and colleagues (2007), participants were asked to: read the question aloud, describe what the question was asking, express how they would respond to the item, and explain why that answer was chosen. Further probing was then used to ask participants to discuss how they interpreted key concepts presented in the measure (i.e., using content-area reading strategy instruction, motivating students to read, and teaching struggling readers; Appendix F).

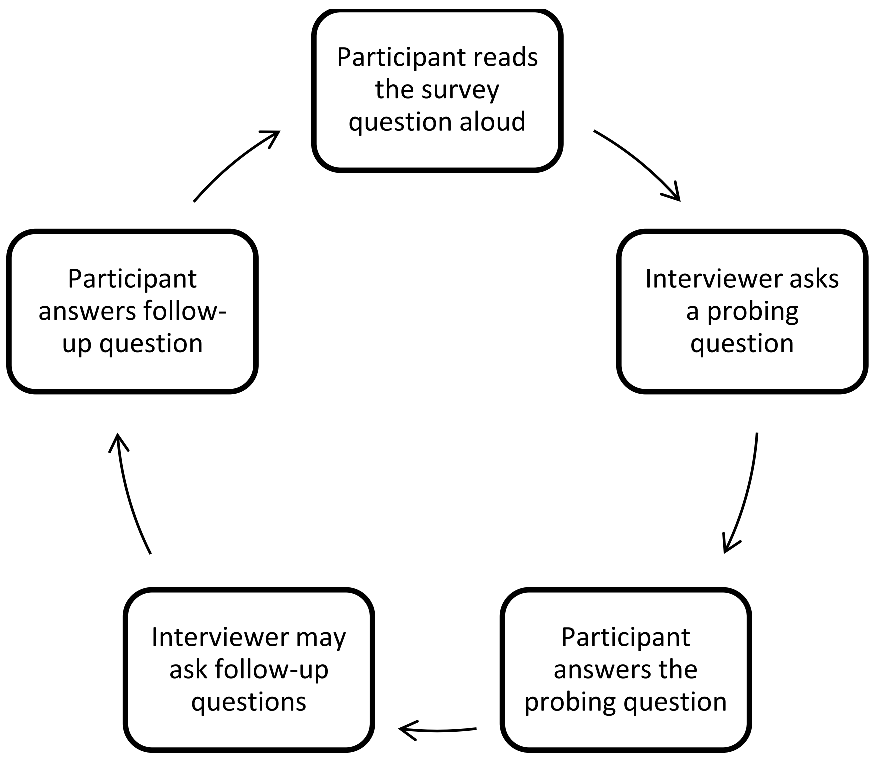

Figure 3. Concurrent verbal probing cycle.

There are four types of probes researchers can utilize when engaging in cognitive interviewing: anticipated (scripted in advance), spontaneous (flexible and derived from the interviewer), conditional (flexible and based on participant behavior), and emergent (flexible and based on participant response; Beatty & Willis, 2007; Conrad & Blair, 2004; Willis, 2005). Both anticipated and conditional probes are pre-determined and scripted in advance by the researcher, with spontaneous and emergent probes being flexible and not scripted in advance (Appendix F). This study used all four types of probes in an effort to glean as much information as possible about the measure from the participants.

Initial cognitive interviews**.** Initial cognitive interviews occurred with secondary content area teachers to determine how members from the target population processed and responded to the 31 items on the STERI instrument. Additionally, participants were asked to answer demographic questions to ensure that the demographic survey to be used in the online survey package was clear.

*Recruitment****.*** Participants were recruited for participation in the study by doctoral students and teachers with whom I was familiar. Teachers were chosen from various high schools using purposeful criterion sampling, thus participants were chosen based upon specific criterion (Patton, 1990). Participants were emailed and asked if they would be willing to be interviewed regarding the clarity of the STERI. Additionally, participants were told interviews would occur in-person at the location of choice by the participant. Seven of the participants were interviewed in their classrooms before or after school, and one was interviewed at a local library after school hours.

Within each content area (e.g., math) a novice (i.e., less than 5 years of experience) and a veteran (i.e., more than 5 years of experience) teacher was chosen. This purposeful sampling occurred for each content area; in sum, there were four teachers each with five years of experience or less, and four teachers each with more than five years of experience. The goal of this purposeful sampling was to get teachers with different years of experience, gender, and cultural backgrounds to ensure that participants with various experiences and perspectives were interviewed. In order to describe the interview participants I collected the following demographic data: primary teaching assignment, years of teaching experience in that content area, years of total teaching experience, gender, cultural background, literacy courses taken for academic credit, number of literacy professional development courses, and highest level of education. In addition, I asked participants to provide some classroom context by answering questions about the rough percentage of students whom were: above-grade learners, on-grade learners, below-grade learners, ELLs, holding an IEPs, and whether teachers worked in a Title I school.

Sample. The sample consisted of eight public high school teachers, three male and five female, from five different schools across the Northeast United States (Table 6). Of the eight teachers, six were Caucasian and two were African American. Teachers had 1-22 years of experience; the average for novice teachers was 2.75 (SD = 1.23) years of experience and the average for veteran teachers was 15.5 (SD = 6.45) years of experience. Most of the teachers held Bachelor's degrees (n = 5), with one having a Masters degrees, and two with a Masters plus additional coursework. Of the eight teachers, only the English teachers reported having more than one literacy course (n = 2) in their coursework, four teachers reported having one course, and two teachers could not remember if they had taken a literacy course. Finally, all of the teachers reported having five or less sessions of professional development dedicated to literacy instruction (M = 3, SD = 1.60).

In addition to teacher demographic data, teachers also responded to questions about classroom context (Table 7). Two of the teachers worked at the same Title I school, all other teachers (n = 6) did not work in Title I schools. Most teachers (n = 5) responded that their classroom had 0-25% ELLs, two teachers had 25-50% ELLs, and one teacher reported 51-75% ELL students. In regards to students with IEPs, two teachers reported this was 0-25% of their students, four teachers responded this was 25-50% of their students, and two teachers stated this was 51-75% of their students.

Additionally, when asked about the percent of students in their classes that were above-grade students, all eight teachers reported this was 0-25% of their classes. The majority of teacher (n = 6) responded that 25-50% of their students were on-grade learners, with the other two teachers reporting a larger percentage of 51-75%. Finally, most teachers (n = 5) stated that 25-50% of their students were below-grade learners, with two reporting 0-25%, and one responding with 51-75%. Thus, overall the majority of teachers (n = 6) taught on-grade students and all of the teachers (n = 8) did not teach more than one-quarter of students that were above-grade learners.

Table 6

Cognitive Interview Participants Self-Reported Demographic Data

	STERI Form	Subject	Years Exp.	Gender	Ethnicity	Ed. Level	# Lit. Courses	Prof. Dev.
Teacher A	A	Science	11	Male	Caucasian	Bachelors	--	3
Teacher B	B	Science	3	Female	Caucasian	Bachelors	1	1
Teacher C	B	English	4	Female	Caucasian	Bachelors+	2	5
Teacher D	A	English	22	Female	Caucasian	Masters+	2	5
Teacher E	B	Mathematics	20	Female	African American	Masters+	--	4
Teacher F	A	Mathematics	1	Male	Caucasian	Masters	1	1
Teacher G	B	Social Studies	9	Female	Caucasian	Bachelors+	1	2
Teacher H	A	Social Studies	3	Male	African American	Bachelors	1	3

Note. Exp. = Experience. Ed. = Education. # = Number. Lit. = Literacy. Prof. = Professional. Dev. = Development. -- = Item left blank.

Table 7

Cognitive Interview Participants Self-Reported Classroom Demographic Data

	Title I	% ELLs	% IEPs	% Above-Grade	% On-Grade	% Below-Grade
Teacher A	No	0-25	25-50	0-25	51-75	0-25
Teacher B	No	0-25	51-75	0-25	25-50	25-50
Teacher C	No	0-25	51-75	0-25	25-50	25-50
Teacher D	No	0-25	25-50	0-25	25-50	25-50
Teacher E	Yes	51-75	25-50	0-25	25-50	51-75
Teacher F	No	0-25	0-25	0-25	51-75	0-25
Teacher G	No	25-50	0-25	25-50	25-50	25-50
Teacher H	Yes	25-50	25-50	0-25	25-50	25-50

Note. All percentages are teachers rough estimates of the number of students in their classes whom belong to each category. ELLs = English Language Learners. IEPs = Individualized Educational Plan.

Procedures and data analysis. After obtaining informed consent (Appendix G), participants were asked to complete the demographic survey (Appendix H) which I used to determine which form of the STERI (A, B) to give participants. Counterbalancing by years of experience and content area was utilized to ensure that participants of similar backgrounds were not interviewed regarding the same items across content areas (e.g., all veteran teachers responding to the same set of items). The rationale for this decision was two-fold: a) having teachers respond to half of the survey is less cognitively demanding and takes less time and b) splitting each factor in half gives each participant in a particular content area the opportunity to engage in questions from each factor (i.e., Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers). The other teacher within that content area was asked the other half of the items from each of the three hypothesized factors.

This counterbalancing was achieved by first ensuring that the first set of teachers from the same content area received different forms of the STERI. After these two interviews, I inputted the STERI form given, content area, and years of experience into Table 6. Then I ensured that the next set of interviews counterbalanced the first (i.e., making sure that the veteran teacher was given the opposite form as the first veteran teacher). Therefore, by the completion of the interviews four teachers (one from each content area), two with more than five years of experience and two with less than five years of experience were interviewed regarding Form A (Appendix I). This same process occurred for Form B (Appendix J). Although the initial goal was to have a counterbalance by gender and cultural background, it became too difficult based on years of experience and content area of the sample; therefore, counterbalancing only occurred by years of experience and content area.

After participants completed the demographic survey, they were asked if they agreed to having the interview recorded, and all teacher (n = 8) consented. Participants were asked questions relating to their understanding of items based on the four probing methods of cognitive interviewing. Although four types of probes exist within the verbal probing method (i.e., anticipated, spontaneous, conditional, and emergent), not all probes were needed for every question across all participants. All participants were asked seven anticipated probes (i.e., four initial probes were asked for every item, and three further probes were asked of each participant; Appendix F), and various follow-up probes. Therefore, each interview included anticipated and spontaneous probes; however, the number of conditional and emergent probes differed across participants based on responses and behaviors.

As the participants responded to items, I took notes in order to aid in the creation of emergent probes, as these were not written in advance. Each interview lasted between 15-30 minutes and all interviews were audio-recorded for analytical purposes. No identifying information was recorded during the interview. Following each interview, I listened to the audio-recorded interviews and took notes in Microsoft Word about participants' responses.

Once all audio recordings were examined, I looked across participants' responses to identify commonalities. These commonalities emerged based on the frequency of responses given by the teachers; therefore, items were revised/removed when multiple participants discussed difficulties with the same items. As the cognitive interviews provided insights into how the items were being read, interpreted, and responded to it was important to look for patterns across participants before dropping items as individual differences are inherent when responding to survey items (Willis, 2005).

***Second round of cognitive interviews*.** After revisions were made to the STERI following the first round of cognitive interviews, I conducted a second round of cognitive interviews using the newly revised instrument (Form C; Appendix K). Interviews followed the same procedure outlined above. Participants were four high school content area teachers (i.e., one from science, social studies, mathematics, and English). Teachers years of experience ranged from 2-9 years (M = 5.25) and two teachers were female and the other two male. Further, teachers described their classroom environment, with most teachers teaching on-grade level students. Finally, one teacher taught in a Title I school. Full demographic information for teachers is available in Table 8, and school information in Table 9.

Table 8

Second Round of Cognitive Interview Participants Self-Reported Demographic Data

	Subject	Years Exp.	Gender	Ethnicity	Ed. Level	# Lit. Courses	Prof. Dev.
Teacher I	Mathematics	2	Male	African American	Bachelors	1	0
Teacher J	Social Studies	7	Female	Caucasian	Masters	0	3
Teacher K	English	9	Female	Caucasian	Masters+	0	9
Teacher L	Science	3	Male	Caucasian	Masters+	1	2

Note. Exp. = Experience. Ed. = Education. # = Number. Lit. = Literacy. Prof. = Professional. Dev. = Development.

Table 9

Second Round of Cognitive Interview Participants Self-Reported Classroom Demographic Data

	Title I	% ELLs	% IEPs	% Above-Grade	% On-Grade	% Below-Grade
Teacher I	Yes	25-50	25-50	0-25	25-50	50-75
Teacher J	No	0-25	0-25	25-50	25-50	0-25
Teacher K	No	0-25	25-50	0-25	25-50	25-50
Teacher L	No	25-50	0-25	0-25	25-50	25-50

Note. All percentages are teachers' rough estimates of the number of students in their classes whom belong to each category. ELLs = English Language Learners. IEPs = Individualized Educational Plan.

Exploratory Factor Analysis

Factor analysis is the most commonly utilized statistical technique to establish construct validation of instrumentation (Osborne & Costello, 2005) and determines how many latent variables underlie a set of items (DeVellis, 2003). Factor analysis is a two-stage process in which researchers first explore and then confirm the factor structure of a particular instrument. The first stage of factor analysis consists of examining the interrelatedness among the items on the scale, called Exploratory Factor Analysis (EFA; Froman, 2001). The second stage of factor analysis tests a specific hypothesis the researcher has about the numeric values of the parameters, defined as Confirmatory Factor Analysis (CFA; Crocker & Algina, 1986). For the purposes of this study, only EFA was discussed in detail as the study is primarily focused on the exploration of the hypothesized three-factor structure of the STERI.

EFA is used during instrument development as a means to: determine how groups of items are related to specific constructs, condense information so that the variance can be accounted for with the smallest number of variables, and define the meaning of the factors or latent variables (Comrey, 1988; DeVellis, 2003). EFA is a technique that uses statistics in relation to correlation, variance, and matrix algebra, but also uses subjectivity in relation to the number of factors identified, the dropping of items, rotational schemes chosen, and the interpretation of factors (Comfrey & Lee, 1992). In this study, EFA was used to determine: the underlying factor structure of the instrument, the amount of variance explained by said factors, and the correlation between the factors in an attempt to meaningfully define a construct. For the purposes of this study, a seven-step process was used to accurately utilize EFA as a method of data analysis (Figure 4).

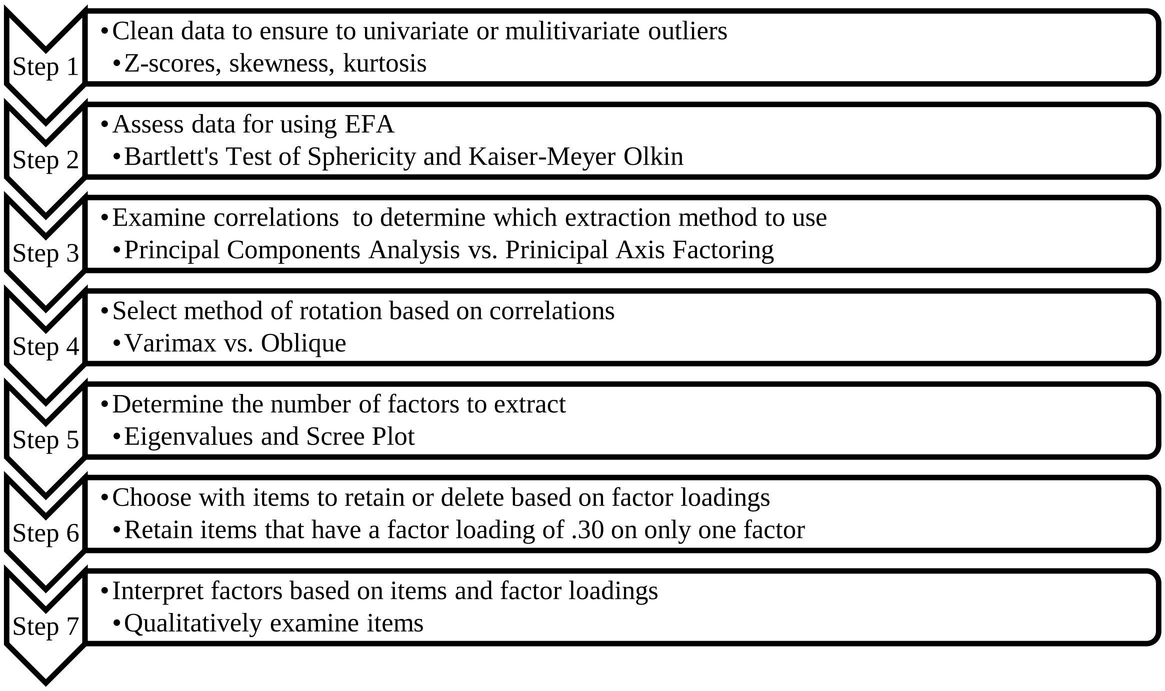

Figure 4. Steps utilized for EFA procedures.

Step One

The purpose of step one was to accurately prepare data to ensure all the assumptions of EFA were met. The assumptions of EFA are comparable to correlational analysis and include: items correlate to at least one other variable, interval level data, linearity, normality, and adequate sample size (DeVellis, 2003). Additionally, as with correlational research, EFA is sensitive to both univariate and multivariate outliers as outliers can skew the findings of the data (Dimitrov, 2012). Two types of outliers can occur in factor analysis data: a) a case as an outlier (i.e., a data entry mistake) or b) a variable or item as an outlier (i.e., an item loads onto its own single-item factor; Osborne, 2014).

First, to account for cases as outliers and to check for any data entry mistakes: standardized z-scores were calculated, histograms and boxplot were examined, frequencies for each item were calculated, and a visual inspection of the responses was conducted. Following these analyses, case outliers were correctly entered into SPSS or deleted from the analysis. Second, when variables act as outliers, the specific variable was deleted and the analyses were re-conducted to examine the remaining variable structure. Additionally, in order to not skew the results of the exploratory analyses, cases were deleted if a substantial (i.e., more than one-quarter) portion of the total survey data reported by a participant were missing or incomplete. By cleaning data and deleting missing or incomplete cases, researchers are able to more adequately conduct an EFA without the potential for bias or nonreplicable results.

Sample size. Another aspect of factor analysis that affects analysis is sample size, which plays an integral role in the integrity of the results presented as well as the overall reliability of the data. As factor analysis is a statistical technique that fits the model to the data, the sample size should reflect a large enough variance that "overfitting" does not occur (i.e., exaggerating the differences of the current sample). Overfitting can result in what Kline (2011) coined the naming fallacy wherein, "just because a factor is named does not mean that the hypothetical construct is understood or even correctly labeled" (p. 230) and can occur more frequently when smaller samples are used (Osborne, 2014). Because factor analysis deals with measurement invariance (i.e., equal factor loadings across groups), small samples limit the statistical power needed to detect a lack of measurement invariance; thus, small samples can sometimes mislead the researcher into thinking the factor structure is stable when it is not (Kline, 2011).

Sample size plays a pivotal role in EFA; however researchers are unable to agree on what constitutes an adequate sample size. Comfrey and Lee (1992) suggested, "the adequacy of sample size might be evaluated very roughly on the following scale: 50 – very poor; 100 – poor; 200 – fair; 300 – good; 500 – very good; 1000 or more – excellent" (p. 217). Others recommended using a minimum of 100 subjects regardless of the number of items (Gorsuch, 1983), to several hundred participants (Cureton & D'Agostino, 1983), or a ratio of 10:1 for sample size to number of variables (Everitt, 1975). For the purposes of this study, the target sample was intended to address the 10:1 ratio provided by Everitt (1975), thus the sample size was based upon the number of items retained after expert review (e.g., if 40 items were retained, a sample of 400 would be necessary). For this study, a sample of 310 was adequate based on the STERI having 31 items, but 400 participants was deemed a more desirable sample.

Step Two

In step two, two tests were used to determine the suitability of the data for an EFA, Bartlett's test of Sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (Kaiser, 1974). According to Dimitrov (2005), Bartlett's test of Sphericity (Bartlett, 1950) should reach a significance of $p < .05$ and the KMO should be greater than .60 for data to be considered appropriate for EFA. Once all survey data were collected and cleaned, step two of the process involved examining the results of both tests to determine if EFA was an accurate statistical test.

Step Three

After Bartlett's test of Sphericity and KMO were met, EFA progressed into step three which entailed determining the extraction method to use. Although there are a multitude of extraction methods (e.g., SPSS has six), for the purposes of the exploratory nature of this study, only Principal Components Analysis (PCA) and Principal Axis Factoring (PAF) were considered. One difference between the two extraction methods is that PCA is used when factors are orthogonal or non-related and PAF is acceptable when factors are correlated with one another. Additionally, PCA assumes that the factors account for all the variance, signifying the expectation of perfect measurement; whereas PAF does not assume that the factors account for 100% of the variance, signifying the expectation of imperfect measurement (Osborne, 2014). Another difference in PCA and PAF extraction techniques is the reported commonalities (i.e., variance accounted for) of items within the scale. As PCA assumes 100% of the variance is accounted for, the initial commonality for each item is 1; but in PAF the initial commonalities represent the amount of variance in each variable is explained by all factors (Osborne, 2014).

This study used both PCA and PAF techniques initially and then I revisited what extraction method to use based on factor correlations. Based upon the adolescent literacy research and the overlapping nature of the factors, I hypothesized the factors were weakly to moderately positively correlated, and therefore PAF would be the extraction method used. In addition, I did not assume that the STERI would provide perfect measurement without error, and therefore PAF was the hypothesized method of extraction for this exploratory analysis.

Step Four

Step four of the EFA process involves the selection of the method of rotation. Rotation was introduced as a way to more easily and reliably interpret the factor structure from the data (Cattell, 1978; Thurstone, 1947). As the axes in factor analysis are indiscriminate, researchers can rotate these axes to make the results of the factor analysis more interpretable without changing the results (Thompson, 2004). There are two common types of rotational methods: orthogonal or oblique rotations. Orthogonal rotations are used when factors are uncorrelated and oblique rotations are used when factors are correlated. Examples of orthogonal rotations are: Quartimax (Carroll, 1953), Varimax (Kaiser, 1958; most common method of rotation and the default in SPSS), and Equimax. Oblique rotations include: Promax (recommended by Thompson, 2004 as the preferred oblique rotation) and Direct Oblimin. Both types of rotational methods (i.e., orthogonal, Varimax and oblique, Promax) were utilized during the initial examination of the data and the subsequent factor structure. Once the correlational analysis between the factors was conducted, the most appropriate method of rotation was used. As stated above, I hypothesized the factors were correlated and therefore a Promax (oblique) rotation would most likely be used.

Step Five

To determine the number of factors to extract, Kaiser's criterion in relation to eigenvalues (Kaiser, 1960) and Cattell's scree plot (Cattell, 1966) were used. Using the adolescent literacy framework and the three hypothesized factors of the STERI (i.e., sense of efficacy for Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers), initial examination of the extracted factors was based on theory. Then, I examined Kaiser's criterion (1960) in which factors are retained when they have eigenvalues of greater than one. According to Kaiser, eigenvalues are the amount of variance that a factor accounts for, and the larger the eigenvalue the more variance a factor explains. One issue with only using eigenvalues is that studies have found mixed results with some research finding an over-extraction of factors (Zwick & Velicer, 1986) whereas others have found an under-extraction of factors (Cliff, 1970). Therefore, an additional source to determine the number of factors extracted was based on the visual examination of Cattell's scree plot. The scree plot visually places factors as points on a plane, with the highest eigenvalues depicted vertically and the lower eigenvalues placed as "scree" or horizontally on the graph. Using Cattell's scree plot, factors are retained that lie above the "scree" on the graph (1966).

In this study, both the Kaiser criterion (1960) of using eigenvalues and Cattell's scree plot (1966) were used examined to determine the number of factors to extract. Using theory, the Kaiser criterion for eigenvalues, and Cattell's scree plot has been shown in research to be an effective way to extract factors during an exploratory analysis (Osborne, 2014). Additionally, I examined the factor loadings and the theoretical meaning of each factor when determining the number of factors to extract; as no method of extraction can provide one-hundred percent accurate results without researcher interpretation.

Step Six

In step six, factor loadings and communalities were examined to determine which items load onto which factors. Factor loadings describe the strength of the relationship between the item and the factor(s). In factor analysis, it is important that each factor be represented with a large enough number of items (variables) so that interpretation of factors can occur. When discussing factor structure, the minimum requirements for factor loadings are used to help researchers determine the factor structure of an instrument. Even though there is no exact acceptable cut value for a factor loading, a widely used recommendation is that loadings should be at least .30 to be considered meaningfully contributing to the factor (Osborne & Costello, 2005). Additionally, communalities represent the portion of variance in the item that is accounted for by the set of factors. In accordance with research, all communalities should be greater than .50 to ensure at least half of the variance in the item is accounted for by the extracted factors (DeVellis, 2003).

According to these recommendations, only items with factor loadings of at least .30 and communalities greater than .50 were included; all items falling below this cut value were examined to determine the proper action (e.g., item revision versus item deletion). Additionally, items that cross-loaded onto two or more factors were examined to determine why cross-loading occurred (e.g., item wording needs to be revised) as these items were not accurately discriminating between factors. As EFA is based upon variance, items were dropped from the instrument one at a time as the deletion of one item changed the commonalities and factor loadings of all other items. Therefore, after a deletion of one item, the EFA was conducted again, new results interpreted, and final decisions made.

Step Seven

The final phase of EFA was the interpretation and labeling of the factors. Once the number of factors to extract has been decided upon, and the factor loading and commonalities for each item is determined, researchers must determine an adequate interpretation of each of the factors. The labeling of factors is qualitative and can be rather subjective as the researcher is drawing his/her own conclusions about the interpretation of factors. In order to reduce researcher biases, upon final extraction of factors I conferred with two research methodologists to corroborate the interpretation of factors.

Reliability and Validity

To determine if the instrument is providing consistent data results reliability is evaluated. Additionally, in order to assess the accuracy of an instrument, validity must be examined throughout the development and use of the scale. Both reliability and validity are essential components of scale development and provide further evidence that an instrument is correctly measuring the intended construct (Crocker & Algina, 1986; DeVellis, 2003).

Reliability

One component of measure development is reliability which deals with the consistency of test scores, or survey responses (Crocker & Algina, 1986). There are many forms of reliability, but for the purposes of this study only internal consistency reliability was addressed; as this is more appropriate when administering an instrument to a single sample on one occasion. This form of reliability is calculated using the average correlations and homogeneity among items within a single administration of a test (DeVellis, 2003). Thus, researchers determine how correlated the items are within the instrument and report the reliability, most often as Cronbach's alpha (α; Cronbach, 1951). Reliability is generally reported for each factor within an instrument and then for the overall instrument itself with .70-.90 considered good and .90 or higher as excellent (Cortina, 1993; DeVellis, 2003).

All quantitative measures have flaws as there is error in all forms of measurement, which is expressed as observed score = true score $\pm$ error score (Crocker & Algina, 1986; Cronbach, Rajaratnam, & Gleser, 1963). The less error in scores the more reliable or consistent the instrument. For an instrument to have perfect consistency the reliability would be equal to 1; therefore, the formula for reliability is 1 – the proportion of error variance to total variance (Kerlinger, 1973). For the purpose of this study, the reliability of the STERI was examined using internal consistency reliability (i.e., Cronbach's alpha) which examined the inter-item correlations for quantitative instruments (Cronbach, 1951). Additionally, the internal consistency implies that items within a scale are homogenous, and thus have a strong relationship to the latent construct under study

(DeVellis, 2003). To examine the internal consistency reliability of the STERI, Cronbach's alpha was reported for each of the three factors (i.e., Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers).

If Cronbach's alpha was between .70 - .90 the subscales were considered to have good internal consistency reliability; with an alpha of .90 or above considered to be excellent (Cortina, 1993; DeVellis, 2003). Thus, if the three hypothesized factors of the STERI (i.e., self-efficacy for Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers) had high internal consistency (i.e., .70 or higher), it would mean that each factor had a specific set of items that all measured that particular factor similarly. As sample characteristics and sample size can greatly influence internal consistency reliability; this study sought to get the largest sample size possible (through ongoing recruitment efforts) across the four content areas equally.

Descriptive Statistics of the STERI

Teacher self-efficacy research has demonstrated that teacher and environmental factors can contribute to teachers' sense of efficacy. Therefore, after conducting an EFA and establishing reliability, various teacher and environmental variables were examined to determine if they had a relationship to participants' responses on STERI items. To examine if contextual variables were related to teachers' sense of efficacy on the STERI measure, various analyses were conducted based upon the type of variables included in the model. First, Pearson's r was calculated for the STERI and all of the continuous contextual variables. Next, ANOVAs, independent *t* tests, and chi-square analyses were conducted to examine the relationship between the STERI and all categorical contextual variables. Then, eleven analyses (i.e., one for each of the contextual variables) were run to determine if differences in teachers' self-efficacy were due to contextual variables. For each ANOVA, self-efficacy was the dependent variable and the contextual factor (e.g., gender, level of education, race/ethnicity) was the independent variable. After running the ANOVAs, when significant differences were found, a post-hoc analysis was run to determine where the statistically significant differences occurred. The type of post-hoc analysis run was determined based on whether Levene's Homogeneity of Variance assumption is violated or not. If Levene's Homogeneity of Variance was not violated (i.e., the variances were equal across content areas), the post-hoc analysis conducted was Tukey's Honest Significant Difference; however, if the variances were not equal across groups the Games-Howell post-hoc was utilized. All differences were reported and explained in Chapter 4.

Hypothesized Correlations Amongst STERI Subscales

I first hypothesized how the three subscales of the STERI (i.e., Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers) were related to one another. This is an important step as it lead to hypotheses about convergent and discriminant validity. In order to examine the relationship between the subscales of the STERI, Pearson's *r* was calculated for each pair of factors (i.e., Factors 1 & 2; Factors 1 & 3; Factors 2 & 3) using SPSS. For the purposes of this study, strengths of relationships were addressed as: a) very weak correlation (r = .20 or less); b) weak correlation (r = .20 - .40); c) moderate correlation (r = .50 - .70); and d) strong correlation (r = .80 and higher). It was hypothesized that all three subscales would be weakly to moderately positively correlated as they all dealt with secondary reading instruction (STERI factors in Table 10). However, the relationships should have only been weak to moderate as each of the subscales represented a different aspect of reading instruction at the secondary level.

Additionally, the strongest positive correlation was hypothesized to be between the factors of Using Content-Area Reading Strategy Instruction and Teaching Struggling Readers, as teachers may have viewed the use of reading strategies as a way to help struggling readers. This relationship should have been positive because these teachers may have associated reading strategy use to helping struggling readers. The Motivating Students to Read factor should have had a weak positive correlation with the Using Content-Area Reading Strategy Instruction factor because teachers may not have viewed reading strategy use as motiving. Further, the Motivating Students to Read factor should have been weakly positively correlated to Teaching Struggling Readers factor, as teachers may have viewed these students as needing motivation, but without drawing a strong connection between the two.

Furthermore, it was hypothesized that these correlations may have changed when the data was disaggregated by contextual variables, more specifically content area. This assumption was first tested when examining the ANOVAs conducted on the contextual variables to determine if significant differences occurred. If none of the ANOVAs showed statistically significant differences across contextual variables, it would be less likely that differences would occur for the correlations based on these factors. However, if differences were shown for the contextual factor of content area, Pearson's *r* was calculated on each pair of the subscales of the STERI for each content area. This analysis tested the assumption that English and social studies teachers may have responded differently to the STERI than science and mathematics. This hypothesis was based on the idea that English and social studies teachers may have more experience and/or knowledge of adolescent literacy instruction than their science and mathematics counterparts.

Table 10

Hypothesized Correlations between STERI Factors and Convergent and Discriminant Scales

	Reading Strategy Instruction	Motivation to Read	Struggling Readers
STERI Factors			
Reading Strat. Ins.	1	+	++
Motivation to Read	+	1	+
Struggling Readers	++	+	1
Convergent Validity			
TSES			
Ins. Strategies	++	0	±
Engagement	0	++	±
GOT			
Mastery Goals	+	++	+
Relational Goals	0	0	++
Discriminant Validity			
Inv. Self & Others			
Self	±	±	±
Others (teachers)	±	±	±

Note. Ins = Instruction. Inv. = Invitations. Reading Strategy Ins. = Content-Area Reading Strategy Instruction. Motivation to Read = Motivating Students to Read. Struggling Readers = Teaching Struggling Readers. ± = Very weakly positively correlated. + = Weakly positively correlated. ++=Moderately positively correlated. 0=No expected correlation.

Validity

When discussing instrument development, validity is a primary concern. Validity is a necessary component of instrument development because it offers support that the instrument is measuring what is purported (DeVellis, 2003). Four areas of validity were examined in this study: face, content, convergent, and discriminant validity.

Face and content validity. Initially, face validity was addressed by having six doctoral literacy students examine the measure and rate items based on clarity and importance to the construct. Further, cognitive interviewing with the target population also provided another point of face validity. Content validity was established by the experts in the fields of self-efficacy and adolescent literacy reviewing the measure.

Convergent and discriminant validity. Convergent validity establishes evidence of similarity of measurement between theoretically related constructs (DeVellis, 2003; Kerlinger, 1973). Hence, convergent validity tests the instrument against another already validated instrument measuring the same construct to ensure that the newly created instrument is accurately measuring what is intended. Discriminant validity tests if an instrument is measuring something it did not intend (DeVellis, 2003). Both convergent and discriminant validity can be assessed using Pearson's *r* to examine the relationship between factors on the newly established instrument to already validated instruments. In this study, to establish convergent and discriminant validity, correlational analysis using Pearson's *r* were conducted for each of the subscales of the STERI (i.e., Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers) and the subscales of the Teacher Sense of Efficacy Scale

(TSES), Goal Orientations for Teaching (GOT), and the Invitations of Self/Others Scale (Table 10).

TSES. The first measure used to assess convergent validity was the two subscales of the TSES: self-efficacy for Instructional Strategies ($\alpha = .86$) and self-efficacy for Student Engagement ($\alpha = .81$; Tschannen-Moran & Woolfolk Hoy, 2001). Consistent with Tschannen-Moran and Woolfolk Hoy's findings, it was hypothesized that the two subscales of the TSES would be moderately to strongly correlated with one another (2001).

Instructional Strategies factor. It was hypothesized that the TSES Instructional Strategies factor would be moderately positively correlated with the STERI Using Content-Area Reading Strategy Instruction factor. This relationship should have occurred as both the TSES and the STERI were measuring teachers' perceived efficacy in the use of instructional strategies. For example, one item from the TSES Instructional Strategies factor was, "How well can you implement alternative strategies in your classroom?" And an item from the STERI Using Content-Area Reading Strategy Instruction factor was, "How confident are you in your ability to teach different reading strategies to students in your content area?" Although the STERI item was more specific and focused on reading strategies, teachers may have responded to both subscales in the same manner due to the use of the term "strategies."

Additionally, it was hypothesized that the STERI Teaching Struggling Readers factor would be weakly positively correlated with the TSES Instructional Strategies factor. This relationship was hypothesized based on close wording of items from the

STERI and TSES. For example, the Instructional Strategies factor had an item that stated, “To what extent can you provide an alternative explanation or example when students are confused?” It was hypothesized that the use of the word “confused” could have caused this factor to be weakly positively correlated with the STERI Teaching Struggling Readers factor as all these items used the term “struggling readers,” which some teachers may have associated with the term confused. Furthermore, there was an assumption that the TSES Instructional Strategies factor and the STERI Motivating Students to Read factor would have had no significant relationship as there was no overarching concepts shared across these two factors.

Student Engagement Factor. I also hypothesized that the TSES Student Engagement factor and the STERI Motivating Students to Read factor would be moderately positively correlated. This relationship was hypothesized to be moderate because both factors assessed and used the term “motivation.” One sample item from the TSES stated, “How much can you assist families in helping their children do well in school?” which is a different approach to engagement than many items on the STERI. An example STERI item for the Motivating Students to Read factor was, “How confident are you in your ability to motivate students to read in your content area?” Therefore, it was hypothesized that the TSES factor of Student Engagement and the STERI Motivating Students to Read factor would be moderately positively correlated as both measure and use the term motivation; however this correlation should have been weaker than the two instructional strategy factors, as the scales approach engagement differently.

Another hypothesis was that the TSES Student Engagement factor and the STERI Teaching Struggling Readers factor would be weakly positively correlated. One item from the Student Engagement scale stated, "How much can you do to motivate students who show low interest in school work?" Some teachers may have associated the phrase "low interest" with "struggling readers," and therefore a weak positive correlation may have occurred between the two factors. Further, there was hypothesized to be no significant relationship between the TSES Student Engagement factor and the STERI Using Content-Area Reading Strategy Instruction factor, as there was no theoretical link between these two factors.

GOT. The second scale used to establish convergent validity was the GOT (Butler, 2007). In 2007 Butler created the GOT scale to demonstrate how achievement goal theory used for learning could also be applied to teaching. Her rationale was grounded in the idea that how teachers approached their own teaching would influence the classroom environment which would impact students in the classroom. To test this hypothesis, she created a four-factor scale to assess the four types of goal orientations teachers have towards teaching: mastery, ability-approach, ability-avoid, and work-avoidance (Butler, 2007).

Butler (2007) described the differences between these four factors as: a) mastery orientation, a teacher striving to learn and continuously develop professional knowledge and skills; b) ability approach, a teacher wanting to demonstrate superior teaching ability; c) ability avoidance, a teacher avoiding the demonstration of substandard teaching; and d) work avoidance, a teacher wanting to make it through the day with little effort put forth.

These four factors demonstrated the four different ways teachers approached their own teaching, not the outcomes of their students; a key distinction from the literature on the goal setting teachers provide for their students. In 2012, Butler added a fifth goal orientation towards teaching that she titled, relational. The Relational Goal Orientation towards teaching measured teachers wanting to create and maintain close and caring relationships with students (Butler, 2012). For the purposes of this study, only the Mastery Goals ($\alpha = .74$) and Relational Goals ($\alpha = .86$) subscales were discussed in detail, as these are the scales being utilized for data analysis. Additionally, I expected the Mastery Goals and Relational Goals subscales to be weakly positively correlated with one another in agreement with Butler and Shibaz's (2014) findings.

When examining the theoretical construction behind the GOT, it became apparent that perceived self-efficacy may play a role in teachers' goal orientations towards teaching. Teachers who have a more mastery goal oriented approach to teaching should also have higher levels of self-efficacy; as they continue to improve their craft, their confidence in their abilities should also increase. Although this relationship has been demonstrated for students (see Bell & Kozlowski, 2002 for an example), it has not yet been explicitly examined with teachers. However, this relationship should occur as a theoretical overlap exists between a mastery goals approach to teaching, higher self-efficacy, and best instructional practices.

Moreover, teachers who have a higher sense of efficacy in performing a specific task should also have higher relational goals for students in relation to that task. Teachers who have high self-efficacy also take responsibility for helping student with special needs and fostering an environment that encourages students' autonomy (Allinder, 1994; Guskey, 1988) which can be associated with teachers wanting to connect with their students and having a relational goal orientation. Therefore, teachers that have a relational goal approach to teaching would care more about forming relationships with their students, much like teachers with higher self-efficacy promoting learning environments that foster similar outcomes.

Mastery Goal Orientation factor. Mastery goal orientations for teaching measures the extent to which a teacher approaches teaching through a lens of continual learning and growth. Teachers who are more mastery goal oriented report higher rates of help seeking behaviors, have positive perceptions of professional development, and have higher intrinsic motivation (Butler, 2007; Retelsdorf, Butler, Streblow, & Schiefele, 2010). In addition, teachers who pursue mastery goals stated that successful days occurred when they learned something new, saw they were teaching better than previously, someone or something in class made them think, and when they overcame difficulties (Butler & Shibaz, 2008; Retelsdorf et al., 2010). Teachers' mastery goal orientation also predicts effective instruction. Teachers who are mastery goal oriented encourage students to seek out help, provide opportunities for students to ask questions, offer challenging and stimulating tasks that promote critical thinking, and support students' mastery goal orientations (Butler & Shibaz, 2008; Retelsdorf et al., 2010).

It was hypothesized that GOT Mastery Goal factor would be positively moderately correlated with the STERI Motivating Students to Read factor. This relationship should have occurred because mastery goals are associated with evaluating competence relative to task demands or prior outcomes, engaging in challenging tasks, attributing outcomes to effort, and utilizing help in the face of difficulties (Butler, 2007). Therefore, teachers with a mastery goals orientation would approach reading instruction through a motivating lens. Moreover, Butler and Shibaz (2014) found that students of teachers with mastery goals had a higher interest in learning; which provides further support for a positive moderate correlation between the two factors.

Further, it was hypothesized that the Mastery Goals factor would be weakly to moderately positively correlated with the STERI factors of Using Content-Area Reading Strategy Instruction and Teaching Struggling Readers. These relationships were hypothesized because a teacher who has a mastery approach towards teaching would continuously be looking to improve his/her teaching skills; something that aligns with the literature on content-area literacy. Additionally, these types of teachers also have higher rates of help seeking behaviors which in turns provides an environment where students feel comfortable asking questions (Butler, 2007, 2012; Butler & Shibaz, 2014). Therefore, there should have been a weak to moderate relationship between the GOT Mastery Goals factor and the STERI factor Teaching Struggling Readers; as teachers who are constantly seeking help may view struggling readers a challenge and way to improve their instruction. Finally, teachers with a mastery approach to teaching like to be challenged and the STERI factors of Using Content-Area Reading Strategy Instruction and Teaching Struggling Readers represent such challenges in teaching, which may cause the factors to be weakly to moderately positively correlated.

Relational Goal Orientation factor. Relational goal orientations towards teaching measures the level to which teachers strive to create and maintain close and meaningful relationships with their students. Teachers who are more relational goal oriented have higher levels of teacher support, teacher involvement, and socioemotional support (Butler, 2012). Therefore, teachers who approach teaching through a relational goal orientation would work to foster relationships with students, take time when they have issues, and listen and get to know their students (Butler, 2012). By creating these meaningful relationships, it has been shown that relational not mastery goal orientations statistically significantly predict students help seeking behaviors (Butler & Shibaz, 2014). Therefore, teachers who take the time to get to know their students create an environment in which students feel comfortable seeking help in the form of questions and clarifications.

It was hypothesized that the GOT Relational Goal factor would be positively moderately correlated with the STERI Teaching Struggling Readers factor. As the Relational Goals factor measured the level of caring and interest a teacher puts forth when teaching, there should have been a relationship between this construct and the construct of teaching struggling readers. In addition, Butler and Shibaz (2014) found that teachers with higher relational goals strove to create close and meaningful relationships with students, which encouraged students to seek out help when they encountered difficulties in school. Therefore, teachers who had higher relational goals supported students' socioemotional well-being and connected with their students in ways that fostered student help-seeking behaviors. Struggling readers are students who are reading at least two grade levels behind and often need extensive remediation. A teacher's relational goals should correlate strongly if he/she forms connections with these types of students. That is, a teacher with higher self-efficacy for helping struggling readers would also have higher relational goals.

Additionally, it was hypothesized that there would be no significant correlation between the GOT Relational Goals factor and the STERI Using Content-Area Reading Strategy Instruction factor as there was no consistent wording or relatable concepts between the two factors. Moreover, relational goals center around the teacher fostering and maintain meaningful relationships with students which has nothing to do with how he/she views content-area literacy instruction (i.e., the hypothesized STERI factor).

Finally, it was hypothesized that the Relational Goals factor and the STERI factor of Motivating Students to Read should have had a weak positive correlation. This relationship should have occurred because teachers may have viewed caring about their students as a way to motivate them to read; however this relationship should not have been strong as there is no overlap in wording or theoretical framework. This relationship would have occurred based on how teachers viewed their relational goals in conjunction with motivating students to read; but the two may not be synonymous for all teachers.

Invitations to Self/Others. The Invitations to Self/Others scale was used to establish discriminant validity for the STERI. Invitations are the messages that one sends to oneself and to others. According to positive psychologists, these messages or invitations can influence their beliefs and behaviors (Purkey, 2000; Purkey & Novak, 1996). Although invitations share many of the same features as social persuasions, one of the four sources of self-efficacy Bandura (1997) hypothesized, they are distinctly different in that they are messages sent rather than received (Morris, 2010; Usher & Pajares, 2006). This distinction has been shown for students when invitations of self/others were statistically significant predictors of self-efficacy when all four of the hypothesized sources of self-efficacy were controlled (Usher & Pajares, 2006). The Invitations to Self/Others was moderately positively correlated with self-efficacy (r = .42 - .49; Usher & Pajares, 2006). Thus, invitations of self/others has been shown to be theoretically distinct from the four sources of self-efficacy in relation to students.

The examination of Invitations to Self/Others and teacher self-efficacy, as measured by the TSES, was conducted in Morris' (2010) dissertation. In his study, he found a non-significant weak positive correlation (r = .14) between invitations teachers send to themselves and teachers' sense of efficacy, and a significant but weak relationship between the invitations teachers send to colleagues and teacher self-efficacy (r = .27; Morris, 2010). These weak correlations established a weak relationship between the invitations teachers send to themselves and others and teacher self-efficacy. This provided additional support that the invitations, or messages, that one sends to themselves and others are theoretically different than perceived sense of efficacy.

Morris' (2010) modified 12-item Invitations to Self and Others Scale was used to test discriminant validity, as it measured something theoretically distinct from self-efficacy. The two subscales invitational messages a teacher sends to himself (α = .89) and the messages one sends to other teachers (α = .73) should have been weakly positively correlated to all three subscales on the STERI; as invitations are distinctly different from perceived self-efficacy (see Usher & Pajares, 2006 for an overview). The Invitations to Self factor should have been distinctly different from the hypothesized factors on the STERI as it involved the types of messages a teacher sent to himself about his teaching in general ("I am quick to recognize my own value as a teacher"; Morris, 2010). Although this item involved the messages a teacher sent to himself regarding teaching, it had no relationship with any of the three hypothesized factors of the STERI (i.e., self-efficacy for: Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers). Additionally, the Invitations to Others subscale should have been weakly positively correlated with the three hypothesized factors of the STERI as this factor encompassed the messages a teacher sent to other teachers ("I like to include others in the activities in my classroom"; Morris, 2010); which was distinctly different than any aspect of teaching self-efficacy.

Therefore, it was hypothesized that the Invitations to Self and Others factors would be weakly positively correlated with all three factors of the STERI. Each subscale should have been weakly positively correlated with the three hypothesized factors of the STERI, with the Invitations to Others subscale having the weakest correlation with all three. This hypothesis was further supported by Morris (2010) who found both the Invitations to Self factor (r = .14) and the Invitations to Others factor (r = .27) weakly positively correlated with general teaching self-efficacy, as measured by the TSES.

Discrimination from the TSES. To provide discriminant validity, as the STERI hypothetically measured the three factors specifically and the TSES generally, it was hypothesized that differences should have arisen when examining data across content areas. That is, when the data were split by content area and each subscale examined, statistically significant differences should have been observed based on content area. It was hypothesized that teachers in English and social studies would have higher self-efficacy on the three factors of the STERI than teachers in math and science; based upon the amount of reading that occurs within each content area.

As additional evidence of discriminate validity, it was hypothesized that the correlations between the three hypothesized factors of the STERI may have been different when examining the data by content area. As stated above, context matters when reporting perceived self-efficacy; thus, content area could have played a role in how teachers responded to items, and effected the relationships between factors. For example, teachers from a particular content area may have more experience in adolescent literacy instruction and therefore could have responded that the three hypothesized factors of the STERI were weakly to moderately correlated; whereas a teacher with less experience may have responded that all three factors were highly correlated. However, when examining the TSES, the correlation among factors should have remained consistent across content areas because the TSES measured general teaching self-efficacy and there should be no reason for a distinction amongst content areas (i.e., all content area teachers should respond in the same manner, therefore the correlations among factors would remain stable).

Procedures and Sampling

After all revisions were made to the STERI based on the ten-step model described above, a sample of high school teachers completed the STERI survey package. Once teachers completed the measure, EFA was conducted to determine the initial factor structure of the measure, internal consistency reliability was assessed, and correlational analyses were examined to establish convergent and discriminant validity (i.e., using the TSES, GOT, and Invitations to Self/Others scales described above).

Instrumentation

The online platform SurveyMonkey (http://www.surveymonkey.com) was used to provide an online version of the STERI to participants. This platform was chosen based on convenience for participants and to maintain anonymity of responses. The overall survey package encompassed six parts (Appendix V).

Part one was informed consent, background information and the purpose of the study (Appendix L). Part two included directions for completing the survey. Part three encompassed demographic questions (Appendix H). Part four was the STERI scale. Part five included scales to assess convergent validity (i.e., TSES, Tschannen-Moran & Woolfolk Hoy, 2001; Goal Orientations for Teaching, Butler, 2007). And part six encompassed a scale to assess discriminant validity (i.e., the Invitations of Self and Others, Morris, 2010).

Demographic survey. The demographic survey included 14 items that addressed various contextual variables related to teaching (e.g., number of literacy PDs attended, years of teaching experience, primary teaching assignment). As stated above, teacher self-efficacy is context, domain, and task-specific and therefore subject to change based upon these factors. According to Bandura's social cognitive theory (1977) a person's behavioral, cognitive, and environmental factors all influence one another reciprocally.

For the purposes of this study I grouped behavioral and cognitive factors into teacher variables, and included all school related variables as environmental variables.

Teacher variables. Research has shown that factors within and among teachers account for differences in self-efficacy beliefs. However, the research has been contradictory in which variables do and do not affect teachers' perceived self-efficacy. The demographic variables I included in the demographic survey were those most associated with teacher self-efficacy throughout the literature.

Gender. Differences in teachers' perceived level of self-efficacy based upon gender have been mixed across the research. Some studies have found that gender does have a statistically significant relationship to teachers' self-efficacy (e.g., Klassen & Chiu, 2010; Raudenbush et al., 1992; Ross et al., 1996). Whereas other studies find that gender is not a significant predictor of teachers' self-efficacy (Skaalvik & Skaalvik, 2007; Tschannen-Moran & Woolfolk Hoy, 2007). One item on the demographic survey addressed gender.

Experience. Years of teaching experience and teacher self-efficacy showed similar patterns as gender across the research. In certain studies (see Klassen & Chiu, 2010 for an example) years of experiences has been shown to relate to self-efficacy, whereas other studies (e.g., Raudenbush et al., 1992; Tschannen-Moran & Johnson, 2011) demonstrate no significance in this relationship. Experience was assessed with two items on the demographic survey, one addressed experience teaching one of the four major content areas (i.e., science, social studies, mathematics, and English) and the other asked for total years of teaching experience.

Level of education and preparation. Studies have not shown a significant contribution of level of education and preparation in predicting self-efficacy (i.e., Raudenbush et al., 1992). However, when studies have examined teachers' level of preparedness (i.e., "I feel well prepared to teach this class"), the results more consistently demonstrate a statistically significant relation to teacher self-efficacy (see Raudenbush et al., 1992 and Ross et al., 1996 for examples).

Moreover, when Tschannen-Moran and Johnson (2011) assessed self-efficacy for literacy instruction through a preparedness lens (i.e., professional development experiences) they found it contributed to teachers' self-efficacy, whereas level of education did not independently contribute to the model. Therefore, both level of education and preparedness items were included in the demographic survey as three separate items. One item asked for teachers' highest level of education and two items asked teachers to report the number of literacy courses taken for credit and the number of literacy professional development sessions attended (i.e., preparedness).

*Race/Ethnicity***.** Race/ethnicity has been shown to contribute to students' sense of self-efficacy (i.e., Pajares & Johnson, 1996; Stevens, Olivarez, & Hamman, 2006), however this relationship has not been reflected in the teacher self-efficacy research (i.e., Raudenbush et al., 1992; Tschannen-Moran & Johnson, 2011). One rationale for the inability to show race/ethnicity as a significant predictor of teacher self-efficacy may be the predominance of white teachers (i.e., 83% in 2011-2012, National Center for Education Statistics, 2013) currently in the field. Although this relationship has not been significant in current research, one item addressed teachers' race/ethnicity on the demographic survey.

Environmental variables. In addition to factors within and amongst teachers contributing to self-efficacy, research has found that certain environmental factors (e.g., level of students) may contribute to one's self-efficacy (i.e., Raudenbush et al., 1992). Three items on the survey were designed to capture school setting: number of courses taught, grade level taught, and Title I status. Based off Raudenbush and colleagues' (1992) findings that track of student (i.e., academic versus non-academic) significantly contributed to teachers' sense of efficacy, five items on the demographic survey addressed the type of students on a more specific level. The purpose for obtaining this information in a more finite manner was to attempt to support the argument that the type of students in a classroom could contribute to teachers' self-efficacy. Two of these questions asked teachers to estimate the percentage of students in their classes who held Individualized Educational Plans (IEPs) and the percentage who were English Language Learners (ELLs). The final three questions asked teachers to approximate the percentage of students in their classes who were: above-grade, on-grade, and below-grade learners. The purpose of asking teachers these questions was to get a sense of the classroom context which the teacher was involved. Moreover, to examine the school environment, one item asked teachers if they worked in a Title I school and another asked for the state in which teachers worked.

Content area. The final contextual factor that may have contributed to teachers' sense of efficacy, particularly in the realm of adolescent literacy instruction, was the teachers' primary content area (i.e., science, mathematics, social studies, and English). First, it is important to note that there was little empirical evidence to support the claim that secondary teachers' self-efficacy for literacy instruction would differ based on content area, as this area has not previously received attention. However, both Raudenbush and colleagues (1992) and Ross and colleagues (1996) demonstrated that high school teaching discipline (i.e., science, social studies, mathematics, English) had an interaction with specific variables that contributed to teacher self-efficacy.

For example, Raudenbush et al. (1992) found that track of students was more impactful for science and mathematics teachers than for social studies and English teachers. They went on to further explicate that when the track of the students was honors, science and mathematics teachers' sense of efficacy was heavily influenced; whereas, this relationship was almost "nonexistent" for social studies and math teachers (Raudenbush et al., 1992). Additionally, Ross and colleagues (1996) found that "feeling successful" (i.e., "I feel successful teaching this class) was a statistically stronger contributor for social studies and English teachers than science and mathematics teachers; conversely, "feeling prepared" (i.e., "I feel well prepared to teach this class") was a statistically stronger contributor for science and mathematics teachers than social studies and English teachers. Therefore, even though both studies used different variables outside of teacher self-efficacy, they both articulated relationships to self-efficacy based on contextual variables (i.e., track of students; feelings of success/preparedness) that differ across content areas. One item on the demographic survey addressed teachers' primary instructional content area.

STERI. The STERI included 31 items and had three hypothesized subscales of perceived self-efficacy: Using Content-Area Reading Strategy Instruction (*n* = 9), Motivating Students to Read (*n* = 10), and Teaching Struggling Readers (*n* = 12). The Using Content-Area Reading Strategy Instruction factor measured teachers' sense of efficacy for using content-area reading strategy instruction in the secondary classroom ("In your content area, how well can you differentiate reading strategy instruction?"). The second factor of Motivating Students to Read, measured teachers' sense of efficacy for using motivational tools to foster a motivational environment for students to read in the content area ("In your content area, how well can you increase students' confidence in reading?"). Finally, the Teaching Struggling Readers factor measured teachers' sense of efficacy to teach and help those students who are classified as struggling readers ("In your content area, how well can you adapt reading materials for struggling readers?"). In addition, the STERI was measured on a 6-point Likert scale of verbal response categories, *Not at all; Very little; Somewhat; Moderately; Quite a bit; and Completely.*

TSES. Eight items from the short version of the TSES (Tschannen-Moran & Woolfolk Hoy, 2001; Appendix M) were used. Included subscales measured teachers' sense of efficacy for: Student Engagement (α = .81) and Instructional Strategies (α = .86; Tschannen-Moran & Woolfolk Hoy, 2001). The Student Engagement factor measured teachers' perceived self-efficacy for engaging students in their classroom ("How much can you do to motivate students who show low interest in school work?"; Tschannen-Moran & Woolfolk Hoy, 2001). The Instructional Strategies factor assessed teachers' perceived self-efficacy for using instructional strategies ("To what extent can you craft good questions for your students?"; Tschannen-Moran & Woolfolk Hoy, 2001). These subscales represented teachers' perceived ability to engage their students in learning and to utilize instructional strategies in the classroom. Each subscale was comprised of four items measured on a 9-point Likert scale ranging from *nothing* to a *great deal*.

GOT. Nine items from the GOT (Butler, 2012; Butler & Shibaz, 2014; Appendix N) were included. The GOT subscales of Mastery Goal Orientations ($\alpha = .74$) and Relational Goal Orientations ($\alpha = .86$) was used (Butler & Shibaz, 2014). The Mastery Goal Orientation for teaching subscale measured teachers' approaches to teaching and whether they focus on building competence in teaching. ("I'd feel I had a successful day in school if I learned something new about myself as a teacher"; Butler, 2012; Butler & Shibaz, 2014). The Relational Goal Orientation subscale measured teachers' motivation for achieving caring relationships with their students and represents a socioemotional component of teaching ("More than anything, I strive to create and maintain meaningful relationships with students"; Butler, 2012; Butler & Shibaz, 2014). In addition, Butler (2012) distinguished the two subscales as "striving to learn" (mastery) and "striving to connect" (relational). Both subscales were measured on a 7-point Likert scale ranging from *very low* to *very high* (Butler, 2012; Butler & Shibaz, 2014).

The Invitations of Self and Others Scale. Twelve items from the Invitations of Self and Others Scale (Morris, 2010; Appendix O), a modified version of the Inviting/Disinviting Index-Revised (Valiante & Pajares, 1999) were used. Invitations are messages that individuals send to themselves and to others that can be either encouraging or discouraging and can influence beliefs and behaviors (Morris, 2010). Morris (2010) adapted the original scale to make the items meaningful for teaching. Therefore, all items related to the self are the types of messages teachers send to themselves (α = .89), and the items related to others encompassed the types of messages teachers send to other teachers (α = .73). A sample item from the *Invitations to Self* subscale was, "I congratulate myself on my own teaching successes" and one from the *Invitations to Others* subscale, "I like to include others in the activities in my classroom" (Morris, 2010). All items were measured on a 9-point Likert scale ranging from *Never* to *Always*.

Procedure

The survey instrument was sent to potential participants via email (see Appendix P), and recruitment occurred from March 26, 2015 to June 17, 2015. The first round of participants were current high school teachers with whom I was familiar. These teachers were contacted through email (personal) and social media (Facebook). Upon completing the survey, these teachers were asked to forward the email and survey on to other high school teachers with whom they are familiar (i.e., snowball sampling). In addition, I contacted secondary teachers across the United States via email through secondary teacher listservs from Teachers.net. These listservs are voluntary and are targeted at the four major content areas (e.g., social-studies-teachers@lists.teachers.net; science-teachers@lists.teachers.net) as a way for teachers to connect with one another across the United States. Emails were sent to each of the four content area listservs every two weeks from March 26, 2015 to June 17, 2015 as a way to promote survey participation. Additionally, to promote participation teachers were given the option of including their email address at the end of the survey to be entered into a drawing to win one of four fifty dollar Target gift cards. These emails were used solely for the purposes of the drawing, and to send the winners the gift cards electronically. Upon completion of the drawing, all email addresses were removed from the data file.

The sample was chosen by self-selection, as the teachers choose to answer the items in the survey package. There was no overlap in participants from any of the previous phases of this study, as all previous participants were recruited via email, and former participants were not sent the survey recruitment email. Additionally, the sample was compared to the current population of high school teachers in accordance with data provided by the 2011-2012 Schools and Staffing Survey to see if the sample was an accurate representation (SASS; National Center for Education Statistics, 2013; Table 11).

Table 11

Demographic Data of High School Teachers from the SASS (National Center for Education Statistics, 2013)

Teacher Demographic	Percentage
Gender	
Male	41.9
Female	58.1
Race/Ethnicity	
White	83
Black	6.2
Hispanic	6.8
Asian	2.0
Pacific Islander	--
American Indian	0.6
2 or more races	1.2
Teaching Experience	
<3 years	9.9
3-9 years	32.7
10-20 years	36.7
>20 years	20.7
Highest Degree Earned	
<Bachelors	4.9
Bachelors	38.2
Masters	47.9
Education Specialist	6.8
Doctor’s Degree	2.1

Note. Variables as listed by the SASS dataset.

All participants were sent the same instrument and asked to answer the items honestly. Participants were not required to answer the survey using any type of identifying information beyond the basic demographic data described above (e.g., years of teaching experiences, gender, race/ethnicity). These data alone cannot be used to identify specific participants. In order to eliminate user error in data entry, all demographic data was entered by participants either checking the box of the answer or with a drop-down menu. This alleviates the limitation of participants incorrectly inputting demographic data (i.e., typing in 222 for years of experience instead of 22).

Participants were unable to view the survey without first clicking a box stating they had read and understood the consent process (part 1). If participants did not consent to taking the survey they were redirected to a page that stated they were unable to participate in the study. If teachers consented to participate in the survey they read through the directions of the survey (part 2). Participants then completed the demographic information (part 3). If participants did not teach one of the four major content areas or did not teach solely high school, they were rerouted during the demographic survey to a page telling them they do not fit the demographics needed to complete the survey. Finally, participants completed the survey and additional instruments for convergent and discriminant validity (parts 4-6). Participants were included in the study if they completed at least half of the entire survey package (Appendix V).

Once the data collection period was over (March 26, 2015 - June 17, 2015), all electronic surveys were exported from SurveyMonkey into an excel spreadsheet. The data were cleaned in excel and then exported into Statistical Package for the Social Sciences® (SPSS®) for Windows, version 20.0.0 (SPSS, 2012). SPSS was used to: analyze all demographic data, conduct EFA, determine internal consistency reliability, and establish convergent and discriminant validity.

Issues with Survey Package

Initially when the survey was created, two items from the demographic portion (i.e., “what is your primary content area?” and “what grade levels do you teach?) were designed to exclude participants who responded to either of these questions with an “other” response; signifying they should not be included in the study and therefore were disqualified. However, after the first week of data collection I noticed that participants were mistakenly being disqualified because they were incorrectly responding “other” to these questions. For example, the survey asked participants to choose their primary content area: “science, social studies, mathematics, English” and some responded “other” and wrote “Physics.” By responding “other” the survey redirected them to a disqualification page; however, these teachers should have been included because physics was meant to be incorporated under the “science” content area.

In addition, although the question regarding grade level taught asked participants to click as many grade levels as applicable, some respondents clicked “other” and typed in “all,” and thus were disqualified from the survey. Therefore, these items were changed to not disqualify participants automatically and the question that addressed teacher’s content area had the responses changed to “English; mathematics; science (any form, e.g., biology, chemistry, physics); social studies/history.” Altering these items helped to not mistakenly disqualify participants; however it also meant that teachers who taught outside the four core content areas and who did not teach in grades 9-12 could potentially complete the survey.

Participant Inclusion

The first step in preparing the data for exploratory analyses was to delete any participants who did not meet the criteria for inclusion into the study. The initial data file contained 478 participants who consented to taking the survey. Of these participants, 43 only completed the demographic portion of the survey and therefore had to be deleted from the data file. Next, 36 participants were deleted because they did not fill out at least half (n = 30) of the survey. Finally, 80 cases were deleted because participants did not teach in a primary content area (n = 59) or did not teach in a high school (n = 21). Upon deleting these cases, the final number of responses that met all criteria for inclusion into the study was 319. Although the goal of data collection was to receive 400 responses (i.e., 100 per content area), obtaining 319 surpassed the bare minimum (n = 310) to meet the 10:1 participant to item ratio needed for EFA suggested by Everitt (1975).

Preparing Data

Next, descriptive analyses were used to check for any data entry mistakes that may have occurred during the exportation process. Standardized z-scores, histograms, means, and standard deviations were examined to determine if any data entry mistakes occurred. If mistakes were present, I inspected each specific mistake in SPSS against the excel spreadsheet and took accurate steps to rectify the issue (e.g., if a 6 was entered in the excel spreadsheet, but it imported into SPSS as a 66; I deleted the extra 6). Next, all demographic data reported from part three of the survey was analyzed using means and standard deviations (continuous data) or frequencies (categorical data). Additionally, distribution tables were created to present this nominal (e.g., content area) or ratio (e.g., years of teaching experience) data. Frequencies were utilized based upon content area taught, in order to describe the sample and discuss possible biases based on self-response rates. These biases could come from an over-response of teachers from one particular content area and/or a lack of response of teachers from one particular content area.

Sample

The sample consisted of 319 teachers of which 209 were female (65.9%) and 108 were male (34.1%). In addition, 246 reported they were white (77.8%), 20 were African American/Black (6.3%), 18 responded as Hispanic/Latino (5.7%), 14 were multiracial (4.4%), 10 identified as Asian (3.2%), 3 reported Hawaiian Native/Pacific Islander (.9%), 2 responded as Middle Eastern (.6%), and there was 1 respondent from each of the following: American Indian, Cajun, and Hellenic (.3% each).

Of the 319 teachers in the sample, 37.9% taught English (n = 121), 23.2% taught science (n = 74), 19.4% taught social studies (n = 62), and 19.4% taught mathematics (n = 62). Teachers across the content areas taught between one and nine classes in their primary teaching assignment, with a mean of 4.73, standard deviation of 1.56, and a mode of 5. Within their primary teaching assignments, 24.7% of teachers taught a single grade level (n = 77) and the majority, 75.3% taught multiple grade levels (n = 235). More specifically, 174 teachers taught grade 9, 150 taught grade 10, 158 taught grade 11, and 130 taught grade 12; note that because teachers taught multiple grades, the total adds up to more than 319. Of the total sample, 31 teachers taught a secondary teaching assignment, 13 of which (41.9%) taught one of the four major content areas, 4 taught special education (12.9%) and the other 14 teachers (45.2%) taught another secondary teaching assignment (e.g., journalism, Spanish, business).

Teachers years of experience teaching in their primary teaching assignment spanned one to thirty-five years, with 58% of the sample having 10 years of experience or less (M = 10.73, SD = 8.1, *Mode* = 1, *Mdn* = 9). Teachers total years of teaching experience ranged from one to forty-one years, with 50.8% having 10 years of experience or less (M = 12.12, SD = 8.5, *Mode* = 1, *Mdn* = 10; See Table 12). The majority of teachers (n = 153, 48.7%) had a master's degree plus additional coursework as their highest level of education. Further, 7% (n = 22) of teachers held a bachelor's degree, 14.3% (n = 45) obtained a bachelors and additional coursework, 24.8% (n = 78) had a master's degree, and 5.1% (n = 16) held a doctoral degree. In addition, 14 teachers also held a specialist degree: 6 had an educational specialist degree, 7 had reading specialist certification, and 1 held a math specialist degree (see Table 13).

Table 12

Years of Experience, Classroom, and Preparation Information for STERI Participants

	Mean	SD	Median	Mode	Range
Years of Experiences in Primary Teaching Assignment	10.73	8.10	9.00	1.00	34
Total Years of Teaching Experience	12.12	8.50	10.00	1.00	40
Number of Classes Teaching	4.73	1.14	5.00	5.00	8
Number of Professional Develop Sessions Attended	6.08	7.70	3.00	0	31
Number of Literacy Courses Taken for Credit	3.74	6.03	2.00	0	31

Note. N = 319.

Table 13

Level of Education for STERI Participants

	Frequency	Percentage
Highest Level of Education		
Bachelors	22	7.0
Bachelors+	45	14.3
Masters	78	24.8
Masters+	153	48.7
Doctorate	16	5.1
Specialist Degree		
Educational Specialist	6	42.9
Reading Specialist	7	50.0
Math Specialist	1	7.1

Teachers were also asked about the number of literacy courses they had taken for academic credit (*M* = 3.74, *SD* = 6.03, *Mode* = 0, *Mdn* = 2): 25.7% (*n* = 80) had not taken a single course, 57.9% (*n* = 180) took between 1-5 courses, 6.7% (*n* = 21) had taken 6-10 courses, 6.7% (*n* = 21) took between 12-20 courses, and 2.9% (*n* = 9) had taken 31 courses. In addition, teachers responded to the number of literacy professional development sessions they attended (*M* = 6.08, *SD* = 7.70, *Mode* = 0, *Mdn* = 3): 18.8% (*n* = 59) had not attended a single session, 49.9% (*n* = 156) attended between 1-5 sessions, 17% (*n* = 53) had taken 6-10 sessions, 9.9% (*n* = 31) took between 11-20 courses, and 4.5% (*n* = 14) had taken 31 courses.

Teachers were recruited from 43 of the 50 states located within the United States, with New Jersey (*n* = 42) and Virginia (*n* = 92) having the highest level of recruitment. The majority of teachers did not teach in Title I schools (*n* = 241, 76%). To examine teaching context teachers were asked to report the percent of students in their classes who were ELLs and those who had IEPs. For both variables, the highest reported percentage was 1-25% of students (ELLs = 208, 65.8%; IEPs = 234, 73.4%); see Table # for full breakdown. In addition, teachers were asked the percentage of students who were above-grade, on-grade, and below-grade readers. For above-grade readers, the majority of teachers (*n* = 174, 54.5%) reported these students made up 1-25% of their students. The majority of teachers (*n* = 180, 56.4%) responded that on-grade readers made up 26-50% of their students. Finally, for below-grade readers the majority of teachers (*n* = 185, 58%) identified that 1-25% of their students were below-grade readers. See table 14 for full breakdown.

Table 14

Teacher Reported Percentages of ELL, IEP, and Reading Levels of Students

	Frequency	Percentage
Percent ELL		
0	48	15.0
1-25%	208	65.2
26-50%	28	8.8
51-75%	13	4.1
75% or more	18	5.6
Percent IEP		
0	12	3.8
1-25%	234	73.4
26-50%	46	14.4
51-75%	11	3.4
75% or more	12	3.8
Percent Above		
0	29	9.1
1-25%	174	54.5
26-50%	63	19.7
51-75%	37	11.6
75% or more	13	4.1
Percent On		
0	7	2.2
1-25%	56	17.6
26-50%	180	56.4
51-75%	56	17.6
75% or more	14	4.4
Percent Below		
0	22	6.9
1-25%	185	58.0
26-50%	73	22.9
51-75%	24	7.5
75% or more	9	2.8

Summary

This study was conducted to develop and validate an instrument to measure secondary teachers' sense of efficacy in reading instruction (STERI) across the four main content areas. The survey was developed using a ten-step approach based off Gehlbach and Brinkworth's (2011) six step approach to survey development. Using methodological suggestions for scale development (e.g., Crocker & Algina, 1986) and previous self-efficacy scale development research (e.g., Klassen et al., 2013) items were created around three hypothesized factors related to secondary literacy instruction (i.e., Using Content-Area Reading Instruction, Motivating Students to Read, and Teaching Struggling Readers).

In addition, expert reviews (n = 10) and cognitive interviews (n = 12) were used as a way to frontload validity before conducting exploratory analyses. Validity is an important component of instrument development as it refers to the degree to which an instrument accurately reflects the construct it is intending to measure (DeVellis, 2003). After items were created two rounds of expert review were conducted to help revise items establish validity before exploratory analyses occurred. To further establish validity, two rounds of cognitive interviews occurred with high school content area teachers. Once all reviews and interviews were analyzed, the STERI was revised and prepared for EFA. After conducting an EFA, Cronbach's Alpha for internal consistency was established. Descriptive statistics on the STERI and teacher and environmental variables were analyzed to determine their contribution to teachers' self-efficacy. Convergent validity was established using the TSES (Tschannen-Moran & Woolfolk Hoy, 2001) and the

GOT (Butler & Shibaz, 2012); whereas discriminant validity was established with the Invitations to Self/Others scales (Morris, 2010).

Chapter Four: Results

The purpose of this chapter is to describe the results of the analyses that were addressed in detail in Chapter 3 in order to develop and validate an instrument to measure secondary content area teachers' sense of efficacy in reading instruction (STERI) across the content areas. First, the results of the second round of expert review are explained. Then, the preliminary cognitive interview data are reported and discussed. Next, the second round of cognitive interview data are presented. Further, the analyses and findings from the Exploratory Factor Analysis (EFA) are described. Then, the reliability results are established. Next, descriptive statistics on the STERI are provided. Finally, the validity results are reported.

Expert Review

Four experts, two from the adolescent literacy field and two from the self-efficacy field reviewed the 31-item STERI. Participants rated each item on the 1-4 scale (i.e., 1 = non acceptable to 4 = exceeds expectations) for both its relevance and clarity; therefore, two scores were reported for each item (i.e., one for relevance and one for clarity; Appendix Q). Additionally, reviewers provided qualitative feedback either on an item-by-item basis, or on the survey overall. In order to analyze the responses from the expert review, all of the responses were inputted into SPSS and Microsoft Word.

As it occurred in the first round of expert review, means were calculated for both the clarity and the relevance of each item, and the qualitative data were analyzed to determine if patterns existed in reviewer comments. Similar to the preliminary expert reviews, revisions/deletions occurred when half (i.e., at least two) of the experts made comments or suggestions to improve item clarity and/or relevance. However, the choice to make changes always favored the field of the experts (e.g., if the content-area literacy experts suggested a change to the literacy portion of an item that was not mentioned by the self-efficacy experts, this was more heavily weighted because of the field of expertise).

Additionally, one of the reviewers offered overall and general feedback instead of item-by-item feedback (see Chapter 3 for details); thus, all tables show only three columns of reviewer qualitative data. Further, some items were problematic for multiple reasons (e.g., low mean score for clarity and redundancy with another item), and therefore some repetition of items within tables occurred. The complete STERI with reviewers' comments can be found in Appendix R. Further, all revisions made to STERI items were first discussed with two educational psychologists, both of whom were familiar with survey development.

Quantitative analysis. Using the previous standards set forth in the preliminary expert review (see Chapter 3 for details), all items obtaining a mean score below 3.00 for clarity or relevance were initially examined to determine if revision or deletion was the best option. Seven of the 31 items had a mean clarity scores of less than 3.00, and one of those item also had a mean relevance score of less than 3.00. Of the seven items with a score below 3.00 for clarity and/or relevance, six items were revised and one was deleted from the STERI. The item that was deleted, “How confident are you in your ability to help struggling readers in your content area?” scored below a 3.00 in both clarity and relevance and was seen as generally “vague” by three of the reviewers. All seven items that scored below a 3.00 for clarity or relevance, reviewer comments on the items, and the final version of the items are presented in Table 15.

Qualitative analysis. After the initial review of items based on the quantitative ratings, a second review occurred based upon the qualitative data received by the expert reviewers. The reviewers offered qualitative notes in three ways: 1) explaining their issues with the item; 2) asking clarifying questions about the item; or 3) offering suggestions for rewording or changing the item. After the qualitative data was analyzed, all 30 of the remaining items in the STERI were revised. The specific types of revisions are discussed in detail below. The complete list of original items, reviewer qualitative notes, and the finalized version of items is located in Appendix R.

Table 15

STERI Items with Mean Scores below 3.00 Based on Expert Ratings of Clarity and Relevance

Original Items	Means		Reviewer Qualitative Notes			Revised Items
	Clarity	Relevance				*"In your content area, how well can you…"*
1. How confident are you in your ability to help students choose the proper reading strategy to improve their comprehension in your content area?	2.67	3.67		Might want to broaden this a bit; "the" proper reading strategy implies there is a correct "one" to choose. Perhaps consider the purpose for which the student would select the strategy?	How well can you help students choose the proper reading comprehension strategy? Or, How certain are you that you can …	*Help students choose which reading strategy to use?*
12. How confident are you in your ability to help struggling readers in your content area?	2.33	2.33	Seems very general—I would let the more specific items carry the day	Too vague; help to do what?	This might not have anything to do with literacy instruction. "To help" is pretty vague.	*Item Deleted*

20. How confident are you in your ability to model the importance of comprehension to struggling readers in your content area?	2.67	3.00		I'm not sure what criteria teachers would use to know whether they had "modeled the importance of comprehension."	I don't think teachers model importance.	*Communicate the importance of comprehension?*
21. How confident are you in your ability to assess struggling readers' difficulties in order to improve their reading in your content area?	2.33	4.00	How is this different from #9 and #13?		Double barreled.	*Help struggling readers improve their comprehension?*
26. How confident are you in your ability to motivate difficult students to read in your content area?	2.67	4.00	Do you mean students who have behavior problems or who have language problems, or who lack vocabulary……?	Define "difficult students"	You might specify what you mean by "difficult students"	*Inspire students of all reading abilities to read?*

29. How confident are you in your ability to motivate students to care about their reading in your classroom?	2.67	3.33		Not sure what this question is asking. See comment for #28.	To motivate them to care? I'm wondering how teachers will understand this. I think it's worth giving some of these items to teachers and doing a think aloud.	*Emphasize the importance of reading?*
30. How confident are you in your ability to use motivational supports to increase students' beliefs about their reading abilities in your content area?	2.33	3.67		What is meant by "increase students' beliefs?" Build their confidence?	What does it mean to increase beliefs? What kind of beliefs? Self-doubt is a belief about one's ability, but I don't think that's what you're after here.	*Increase students' confidence in reading?*

Note. All item numbers reflect the original numbering of the STERI for expert review.

Types of revisions. Based on the data provided by the expert reviewers, revisions occurred to the remaining 30 items on the STERI, and some duplication of items in tables happened. First, the stem of the items was changed to reflect a more social cognitive approach (i.e., focusing on the action of doing, not thoughts about doing). Second, eight items were revised or deleted based upon redundancy with other items. Third, two items were changed to address the issue of being double-barreled; therefore the item was asking two questions instead of one. Fourth, eight items were revised to reflect better word choice and eight items were changed to be more concise. Fifth, in order to not have superficial factor loadings based upon common terminology, key terms associated with each factors were changed within certain items. Finally, two items were added to the measure to replace the core concepts presented by the two deleted items.

Items stem. One of the first revisions that occurred was to change the stem of the items from "How confident are you" to "In your content area, how well can you…" This stem more aligned with Bandura's (1977) conception of self-efficacy as thoughts about one's perceived abilities right now; not thoughts about future abilities. The change of stem came as a suggestion from E. L. Usher who articulated, "Does this measure confidence or perceived capability? This is really nuanced, but there might be a difference in this and in my rating of how well or whether I think I can do something" (personal communication, November 12, 2014). Therefore, the stem of the item was changed to reflect a more complete notion of self-efficacy using Bandura's (1977) original conception. Further, the stem was placed at the top of the measure instead of being included at the beginning of each item. The rationale for this change came from A. W. Hoy who stated, “Having the repeated phrases in each item makes the task seem longer, more daunting, and may strain on working memory” (personal communication, October 28, 2014).

Repetition. As reported by the expert reviewers, four sets of items (i.e., eight items total) were repetitive with one another. Therefore, the wording of items was changed to reflect more nuanced aspects to the items and to decrease the overlap. For example, two items that overlapped were, “In your content area, how well can you diagnose why a student is struggling with reading?” and “In your content area, how well can you diagnose the difficulties of struggling readers?” To differentiate the two items, they were changed to: “In your content area how well can you determine why a student is struggling with reading?” and “In your content area how well can you determine the reading needs of struggling readers?” respectively. Further, one item was deleted “How confident are you in your ability to model effective reading skills to help struggling readers in your content area?” as it could not be adapted to be different enough from the item “How confident are you in your ability to model effective reading in your content area to help struggling readers?” These types of revisions occurred for each of the four sets of overlapping items (See Table 16).

Table 16

Eight STERI Items Revised Based on Redundancy and Overlap

Item	Reviewer Qualitative Notes			Final Items
				"In your content area, how well can you…"
2. How confident are you in your ability to help students choose the proper reading strategy to improve their comprehension in your content area?		Might want to broaden this a bit; "the" proper reading strategy implies there is a correct "one" to choose. Perhaps consider the purpose for which the student would select the strategy?	How well can you help students choose the proper reading comprehension strategy? Or, How certain are you that you can …	*Help students choose which reading strategy to use?*
6. How confident are you in your ability to encourage students to find the reading strategy (or strategies) that work(s) best for them in your content area?	This is a bit close to #2	This one overlaps with #2, but is clearer in that it references the selection of multiple strategies.		*Encourage students to find the reading strategy (or strategies) that work(s) best for them?*
9. How confident are you in your ability to diagnose why a student is struggling with reading in your content area?		The "diagnose" sounds like something within the realm of reading specialists.	See above. Alternative: How certain are you that you can diagnose …	*Determine why a student is having difficulty with reading?*

13. How confident are you in your ability to diagnose the difficulties of struggling readers?	How is this different from #9?	Overlaps with #9		*Determine the reading needs of struggling readers?*
10. How confident are you in your ability to model effective reading in your content area to help struggling readers?			How is this different from #17 below?	*Model effective reading for struggling readers?*
17. How confident are you in your ability to model effective reading skills to help struggling readers in your content area?	How is this different from #10?	Overlaps with #10		*Item Deleted*
14. How confident are you in your ability to differentiate (or change) literacy instruction for struggling readers in your content area?			Choose differentiate or change	*Change your instruction to meet the needs of individual readers?*
18. How confident are you in your ability to adjust reading materials in your content area for struggling readers?	Are you getting at something different from #s 14 and 16?	Overlaps with #14		*Adapt reading materials for struggling readers?*

Note. All item numbers reflect the original numbering of the STERI for expert review.

Double-barreled. Another type of revision that occurred was to change the two items that were double-barreled; thus the items were asking two questions instead of one. To address these issues, items were simplified by ensuring that they were only asking teachers to respond to one specific task. For example, the item "In your content area, how well can you differentiate or modify reading strategies for different levels of readers?" was changed to, "In your content area, how well can you differentiate reading strategy instruction?" Thus, the revised question asked teachers to respond to their ability to change their reading strategy instruction, instead of the initial question which asked their ability to change their instruction to help readers of different levels.

The other double-barreled item was changed from, "In your content area, how well can you assess struggling readers' difficulties in order to improve their comprehension?" to "In your content area, how well can you help struggling readers improve their comprehension?" Again, the item was changed to ask participants only about helping struggling readers improve comprehension, as opposed to the original item which asking about assessing struggling readers difficulties and helping to improve their comprehension. Therefore, the issue of double-barreled items was rectified by revising the two items to ensure only one task was being assessed.

Word choice and conciseness. Two other ways in which items were revised included: changing word choice and making items more concise. Both changes reflected an attempt to clarify issues with wordiness and understanding among these eight items. An example of how word use was changed was revising, "How confident are you in your ability to diagnose why a student is struggling with reading in your content area?" to "In your content area, how well can you determine why a student is having difficulty with reading?" One of the adolescent literacy experts, described how the use of "diagnose" seemed too close to the realm of a reading specialist, thus the wording of the item was changed to reflect this concern (M. E. Mraz, personal communication, November 8, 2014). A full list of the eight items changed due to word choice is in Table 17. One example of item revision to address conciseness was the change from, "How confident are you in your ability to teach different reading strategies to students in your content area?" to "In your content area, how well can you teach different reading strategies?" See Table 18 for a complete list of the eight items changed to reflect more conciseness in language.

Additionally, in order to not have items superficially load together on a factor based on similar word use, the term *motivation* was removed from five items and the term *struggling* or *struggling readers* was removed from four items. Instead of using these specific terms, synonyms were used (e.g., encouraged instead of motivated) or the terms were removed all together. Making these revisions may help to alleviate the issue of items loading together onto a factor due to all items using the same term (i.e., motivation or struggling readers). The final version of the STERI after expert review is located in Appendix S.

Addition of items. The original STERI had 31 items; however one was deleted due to a low mean clarity and relevance. A second item was deleted based on its repetition with other items. Although these items were deleted, reviewers acknowledged the underlying concepts of the items were important. Therefore, two additional items were added to the STERI to represent concepts lost from the deleted items. The two items, “In your content area, how well can you: utilize a variety of texts to encourage students to read; identify improper reading strategy use?” were added to the measure and brought the total number of items on the STERI to 31.

Summary of expert review. After revisions were made to the remaining 29 original and two new STERI items, the 31-item survey was again reviewed with two educational psychologists. Further modifications and revisions were made to increase item clarity. Therefore, some of the item revisions do not have expert review comments included as those items were revised after I consulted with the two educational psychologists. At the end of the expert review the STERI contained 31 items: nine for the Using Content-Area Reading Strategy Instruction factor, 10 for the Motivating Students to Read factor, and 12 for the Teaching Struggling Readers factor. The finalized version of the STERI after expert review can be found in Appendix S.

Table 17

Eight STERI Items Revised Based on Word Choice

Item	**Reviewer Qualitative Notes**		**Final Items**
			"In your content area, how well can you…"
3. How confident are you in your ability to help students use reading strategies to improve comprehension in your content area?		This seems redundant with the previous.	*Help students apply reading strategies to improve comprehension?*
5. How confident are you in your ability to model how an effective reader uses reading strategies in your content area?		How well can you model effective reading strategy use?	*Model effective reading strategy use?*
9. How confident are you in your ability to diagnose why a student is struggling with reading in your content area?	The "diagnose" sounds like something within the realm of reading specialists.	See above. Alternative: How certain are you that you can diagnose …	*Determine why a student is having difficulty with reading?*

14. How confident are you in your ability to differentiate (or change) literacy instruction for struggling readers in your content area?		Choose differentiate or change	*Change your instruction to meet the needs of individual readers?*
16. How confident are you in your ability to differentiate course readings to help struggling readers in your content area?			*Select course readings to match the needs of individual readers?*
19. How confident are you in your ability to help struggling readers assess their comprehension in your content area?	Are you referring to a student's ability to self-assess?		*Help struggling readers self-assess their own comprehension?*
23. How confident are you in your ability to motivate students to read in your content area?			*Encourage students to self-assess their reading comprehension?*
24. How confident are you in your ability to use differentiation as a motivation tool for reading in your content area?		Do teachers know what this means? (I'm not sure I do.)	*Modify instruction to motivate students to read?*

25. How confident are you in your ability to motivate students to improve their comprehension when reading in your content areas?	I don’t think this is particularly clear. To motivate to improve one’s understanding? Maybe …	*Motivate students to improve their understanding when reading?*

Note. All item numbers reflect the original numbering of the STERI for expert review.

Table 18

Eight STERI Items Revised Based on Conciseness

Item	Reviewer Qualitative Notes	Final Items
		"In your content area, how well can you…"
1. How confident are you in your ability to teach different reading strategies to students in your content area?	Omit "to students"	*Teach different reading strategies?*
7. How confident are you in your ability to help students monitor their own use of reading strategies in your content area?		*Help students monitor their use of reading strategies?*
8. How confident are you in your ability to assess students' use of reading strategies in your content area?		*Assess students' use of reading strategies?*
11. How confident are you in your ability to help struggling readers		*Help struggling readers comprehend text?*

comprehend text in your content area?				
15. How confident are you in your ability to create intervention plans to help struggling readers in your content area?				*Create intervention plans to help struggling readers?*
22. How confident are you in your ability to use a variety of assessments to help struggling readers in your content area?		Might consider specifying formal v. informal assessment, or formative v. summative assessments		*Use a variety of assessments to help struggling readers?*
27. How confident are you in your ability to motivate students who are not interested in reading in your content area?	This is clearer.		Hmmmm. Motivate them to do what?	*Encourage uninterested students to read?*
28. How confident are you in your ability to motivate all students to read in your content area?	This may be a bit general	“Motivate all students” is pretty broad. Are you referring to engagement or to applying content area reading outside of the classroom?		*Motivate students to read?*

31. How confident are you in your ability to use various types of assessments to motivate students in reading in your content area?	As with question #22, you may wish to specify certain types of assessments.	*Use various types of assessments to motivate students to read?*

Note. All item numbers reflect the original numbering of the STERI for expert review.

Initial Cognitive Interviews

Cognitive interviews were conducted with eight teachers from the target population (i.e., secondary content area teachers) to determine how they viewed items on the STERI. Cognitive interviewing can aid researchers in understanding how the target audience understands, process, and responds to items on the instrument (Willis, 2005). Therefore, using cognitive interviews aided in formulating better items and revising problematic items before exploratory analyses occurred (see Chapter 3 for complete details on interview process). Further, these interviews established content validity for the instrument as these teachers were from the target population (i.e., secondary content area teachers).

Eight participants engaged in cognitive interviewing to determine if further revisions to the STERI items was needed. Demographic data for these teachers is presented in Table 6 and 7. Four teachers were interviewed regarding Form A (i.e., one from each content area) and four about Form B (i.e., one from each content area) of the STERI. Form A included 16 items and Form B contained the other 15 items of the STERI (see Appendices I & J).

During cognitive interviewing (regardless of the STERI form), participants were encouraged to respond to questions/items in the manner they were most comfortable. This led to each interview differing in how participants responded: talking through the content of the items, putting the items into their own words, describing the items as easy or difficult, and/or responding to how they would answer the items. For example, the male science teacher with 11 years of experience read every item aloud, explained the item in his own words, and referenced how he would respond to the item; whereas the female English teacher with 22 years of experience read each question aloud, then fluctuated between answering the items and explaining each answer or talking through what she believed the item was addressing. Because cognitive interviewing can be cognitively taxing, participants were encouraged to respond to items in the way they were the most comfortable, and conditional probing was used to get participants to expand on their responses when necessary. Even though the participants' process for reviewing the items differed, all cognitive interviews achieved the same goal: to assess the clarity and teachers' understanding of the STERI items.

After listening to the recordings of the cognitive interviews for Forms A and B and analyzing the data, teacher responses were organized into five separate categories to represent different components of their interviews. These categories were based upon any issues participants described when answering items, the ease or difficultly in rewording items, and any problems participants had answering items. In addition, as no two interviews occurred in the exact same manner, it was impossible to exactly align participants' responses for each item. For example, certain items had three of the four participants who demonstrated understanding of the item by putting it in their own words and then explaining how they would answer the item; however, the fourth participant may have demonstrated understanding by saying the item was clear and then putting it into his/her own words. Thus, the information encompassed in each category is similar based on the thematic nature of the participants' response; however the way in which the participants responded may have differed. The five categories items were placed into were those teachers: 1) stated were “easy” to answer, 2) demonstrated understanding in their responses, 3) explained as “difficult” to answer based on implementation, 4) articulated as “difficult” because they did not see it as part of their job, and 5) stated were “unclear” based upon the wording of items. Further, any revisions made to items were first discussed with the aforementioned educational psychologists in the same manner addressed after the expert reviews.

Ease of items. One of the first ways participants demonstrated their understanding of items was when they described items as: “easy,” “clear,” “simple,” and stated, “I understand what the question is asking.” In addition, for the items participants said were “easy” they were able to easily follow-up this comment by restating the item in their own words or responding with how they would answer the item. For example, for the item “In your content area, how well can you adapt reading materials for struggling readers in your content area?” on Form B of the STERI, Teacher G stated, “Yes I completely get the question, it’s asking if I change up what materials I give lower level readers, and I would say I do this somewhat to moderately because I’m always trying to differentiate” (social studies, nine years of experience, female). Here Teacher G was able to articulate that the question was easy to understand, describe how she demonstrates this task, and provide an answer that she would give on the survey.

Across both forms of the STERI nine items (three on Form A and six on Form B) were labeled in a similar pattern by participants (see Table 19). The most likely conclusion for the greater number of “easy” responses on Form B is due to the more concise nature of these six specific items. As the forms were created by randomizing the items in excel and then splitting them into two groups, this randomization is most likely the rationale for the discrepancy across forms. Further, the nine items labeled as "easy" spanned all three subscales equally across both forms (i.e., Form A: one item per subscale, Form B: two items per subscale); therefore, the subscale of the item did not dictate the item's perceived difficulty level.

Table 19

Example Items from STERI Teachers Found "Easy" to Answer

Items	**Content Area of Teacher**			
Form A	Social Studies ***Teacher H*** 3 years of exp. Male	Mathematics ***Teacher F*** 1 year of exp. Female	Science ***Teacher A*** 11 years of exp. Male	English ***Teacher D*** 22 years of exp. Female
1. Model effective reading strategy use?	"This question is very clear…can I model correct reading strategy use."	"I don't do it, but yes I understand what you're asking me here."	"I understand the question, and I would say very little because I often just refer them to the toolbox, I'm not modeling myself very often."	"Yes, you're asking if I think I can effectively model how to use reading strategies."
6. Model effective reading for struggling readers?	Reworded the question: "Can I model good reading for students who don't read well?"	"Like I said before (laughs), I'm not doing a lot with reading but I definitely know what the question is asking.	"It's clear to me what you're asking me, and unfortunately I'd say very little to not at all because again I refer them to the toolbox or try to find something easier for them to read."	Reworded question: "How well do I think I show students who struggle how to read better?"

14. Encourage uninterested students to read?	"I think this is simple and to the point; I don't know how to reword it other than to say, can I get uninterested students to read."	Reworded question: "Can I get students who don't care to read?"	"I understand what you're getting at; here I think I'd answer quite a bit because I'm always trying to find them other science related things to read to try and get them into the content."	"This is very clear; can I get students who don't care about English to read. If I could do this my life would be a lot easier (laughs)."
Form B	Social Studies ***Teacher G*** 9 years of exp. Female	Mathematics ***Teacher E*** 20 years of exp. Female	Science ***Teacher B*** 3 years of exp. Female	English ***Teacher C*** 4 years of exp. Female
1. Teach different reading strategies?	"This is an easily understood question, and even though I don't teach reading strategies daily, I think I could moderately teach them effectively."	Rewording of question: "Do I teach different reading strategies in my classroom?"	"This is a question I understand, but I don't teach reading strategies in my classroom. I offer different types of reading materials, but not strategies explicitly."	"This is very clear and something I do often. This is English, if they can't read it than they're doomed. I don't teach strategies per say but I teach them how to be successful reading and what I do when I get stuck."

2. Emphasize the importance of reading?	"I can't even rephrase this without just saying the same thing again, but yes I emphasize reading quite a bit in my classroom."	Rewording of question: "Can I stress to students how important it is to read?"	"Yes, I tell them all the time how important it is to read in science because labs, tests, and everything is text related. So yes I'm always telling them to read."	"This is super clear. I do this all the time. Again it's English, they have to read because everything comes from the text."
11. Identify improper reading strategy use?	"I don't know that I could necessarily know right away that a student was using a strategy wrong, but I would probably see it in their writing and inability to do well on tests."	Rewording of question: "Do I know if a student is using a reading strategy wrong? Honestly I know I cannot do this, maybe very little if that."	"I don't focus on reading strategies in my classroom because I teach science, so I think I do this very little to not at all."	"This is clear and I understand the question, but since I don't have formal training in reading strategies I don't know that I'd know if a student was using one wrong."

Note. YRS = Years of Experience. G = Gender. Item numbers are from original STERI for cognitive interviews form.

Understanding of items. Participants demonstrated understanding in their ability to: rephrase items into their own words, describe how they would answer items, and/or explain what they believed the item was asking. An example from Form A was in response to the item, “In your content area, how well can you help struggling readers comprehend text in your content area?” where Teacher A said, “Yes, well I’d have to say moderate to quite a bit because I am always pointing them to their toolbox to help them find ways to make meaning” (science, 11 years of experience, male). Teacher A was able to show his understanding of the item by describing how he would answer the item and explaining why he would answer that way. More examples of participants’ demonstrating understanding of items can be found in Table 20.

Across both forms of the STERI, 13 items (five on Form A and eight on Form B) were clear based upon participants’ ability to demonstrate meaning through rewording, answering, explaining, and responding to said items. Additionally, no patterns arose when examining the 13 items across the three subscales, as distribution of items was nearly equivalent across subscales. For Form A, two items were on the Motivating Students to Read subscale and three items were on both the Using Content-Area Reading Strategy Instruction and Teaching Struggling Readers subscales. Similarly, on Form B, two items were on the motivation subscale, two items were on the reading strategies subscale, and one item was on the struggling readers subscale. Thus, the clarity of items based on participant understanding was not related to the subscale of the item.

Table 20

Items from STERI in which Teachers Demonstrated Understanding

Items	**Content Area of Teacher**			
	Social Studies	Mathematics	Science	English
Form A	***Teacher H*** 3 years of exp. Male	***Teacher F*** 1 year of exp. Female	***Teacher A*** 11 years of exp. Male	***Teacher D*** 22 years of exp. Female
5. Change your instruction to meet the needs of individual readers?	Reworded question: “Do I change my instruction to help all my students?”	“Yes, what works for some will definitely not work for others; especially in my special education classes."”	“Because it’s content specific to science so it is extremely important that they read and understand what we’re reading, so I have to mix things up.”	“They [students] all need something a little bit different and so I’m constantly changing up what I’m doing and what I’m giving them to read to help them get the content.”
15. Inspire students of all reading abilities to read?	“Yes, I know what you’re asking with this question…can I get all students to read.”	“Hmmm…I think it’s somewhat to very little [referring to anchor scores], in math I don’t really think about their reading ability.”	“I’d say quite a bit because they have to read to get the grade and so I’m always giving them different types of things about science to read to get them to want to read for understanding.”	“I do this a lot because in English we have no choice, they have to read and so I’m constantly pushing them to want to read no matter the level, they have to read in order to succeed.”

Form B	Social Studies ***Teacher G*** 9 years of exp. Female	Mathematics ***Teacher E*** 20 years of exp. Female	Science ***Teacher B*** 3 years of exp. Female	English ***Teacher C*** 4 years of exp. Female
5. Use a variety of assessments to help struggling readers?	"I would say moderately to quite a bit because I'm always trying to change up my assessments to give students of all abilities to opportunity to do well."	Rewording of question: "You're asking if I change of my instruction for all types of students, even those who struggle reading."	"Yes, I am often changing the types of assessments I'm giving to help all students succeed in science. It's not always a reading issue, more an understanding of the content which could be related to reading or not."	"Yes, and especially when we begin a new topic I'll do things that are related and are more creative to motivate them; so I'll go very personal with some assessments and then work towards a more material driven approach."
6. Help students choose which reading strategy to use?	"This question is very clear, do I help students choose the reading strategy to use (laughs); sorry I didn't mean to rephrase it exactly like it's written but it's clear."	Rewording of question: "Do I help students choose reading strategies?"	"I don't particularly do that with each individual student because I'm not a reading classroom, I may mention a way to help them understand but I'm not helping them choose strategies per say."	"Yes, and I also work with them on choosing different strategies at different times, like sometimes you need to use prediction but other times that won't work. So I try to help them see there's more than one strategy they can use."

Note. YRS = Years of Experience. G = Gender. Item numbers are from original STERI for cognitive interviews form.

Difficulty in implementation of items. Another category of items were those participants stated were "difficult to answer" based on the implementation of the item's task in their content area. These types of items were difficult based on the task described, but not difficult due to the wording or clarity of the items. For example, during each of the interviews using Form A, there were three items that participants struggled to answer and stated things such as, "that one is difficult" or "that's a tough one." To assess if the item was difficult to answer, based on the wording or on the task; the follow-up probe used was, "So do you think the question is difficult because of the wording, or it's asking too many things, or is it something else?" One example of this was Teacher D's response to the item "In your content area, how well can you motivate students to improve their understanding when reading?" She stated, "If I could do that I would be a better teacher," and then responded to the follow up question and articulated, "For me, maybe being a long-time teacher, just the answer itself is difficult…motivating students is just a bear of a topic" (English, 22 years of experience, female). In this case, the teacher emphasized that the answer was difficult because the task being assessed was one she personally found difficult. For each of these three items, participants stated the task itself was inherently difficult, but the wording and clarity of the item was not problematic.

On Form A, three separate items were described as difficult based on their implementation; however only one item was seen as difficult by more than one participant (see Table 21). Additionally, there were no items on Form B that participants labeled as difficult to implement. Of the three difficult items on Form A, each was from one of the three subscales; therefore difficulty in an item's task was not related to the subscale.

Table 21

Items from STERI Teachers Found "Difficult" to Answer Based on Implementation

Items	**Content Area of Teacher**			
	Social Studies	Mathematics	Science	English
Form A	***Teacher H*** 3 years of exp. Male	***Teacher F*** 1 year of exp. Female	***Teacher A*** 11 years of exp. Male	***Teacher D*** 22 years of exp. Female
5. Change your instruction to meet the needs of individual readers?			"I understand the question, but it's a tough question because we have one basic book, so often I'm going online to find other texts."	
7. Assess students' use of reading strategies?	"I know what the question is asking, do I know when they are using a strategy incorrectly, but I'm not even sure how to go about doing that, that's not my forte, I don't feel strong in that area."			

8. Motivate students to improve their understanding when reading?	“I’m not sure any teacher can really do this. I can tell them over and over again to read the textbook but that doesn’t mean they’ll do it. I’m not sure I really have any influence in this area.”		“If I could do that I would be a better teacher” follow up “so do you think the question is difficult because of the wording, or it’s asking too many things, or is it something else” teacher “For me, maybe being a long-time teacher, just the answer itself is difficult…motivating students is just a bear of a topic.”

Note. YRS = Years of Experience. G = Gender. Item numbers are from original STERI for cognitive interviews form.

Difficulty in applicability of items. In addition to participants finding items difficult based on the task delineated by the item, other items were considered difficult because teachers did not see the items as being a part of their job description. For example, in response to the item, “In your content area how well can you determine the reading needs of struggling readers?” Teacher H stated, “To me, that’s a difficult one because I’m in a social studies classroom and so I’m working on the content and this seems to me like something the reading specialist should be doing” (social studies, three years of experience, male). Therefore, Teacher H was not having difficulty with the wording of the item or the items task itself, but instead he saw the task as something outside of his realm of teaching and that made it difficult to respond to the item.

On Form A, three items were considered difficult by the social studies, mathematics, and science teachers based on their beliefs regarding instruction in their content area; whereas no items on Form B had these issues (see Table 22). Although the social studies, mathematics, and science teachers had difficulties with these three items, the English teacher was able to reword these items and expand on the items’ content. Therefore, it was concluded that these items did not need to be altered as the difficulties surrounding the items were based on individual beliefs, not the items themselves. Interestingly, all three items were from the Teaching Struggling Readers subscale and all three teachers who had difficulties described these items as involving: “the reading specialist”; something done in “English”; and as something outside of their job description.

Table 22

Items from STERI Teachers Found "Difficult" to Answer Based on Their Teaching Responsibilities

Items	**Content Area of Teacher**			
Form A	Social Studies ***Teacher H*** 3 years of exp. Male	Mathematics ***Teacher F*** 1 year of exp. Female	Science ***Teacher A*** 11 years of exp. Male	English ***Teacher D*** 22 years of exp. Female
9. Determine the reading needs of struggling readers?	"To me, that's a difficult one because I'm in a social studies classroom and so I'm working on the content and this seems to me like something the reading specialist should be doing."	"That's what a specialist is for" Follow up: "do you use the reading specialist?" "I don't know if we have one anymore."	"I don't know if that really applies to me. I mean they have the strategy toolbox in the front of the book to use if they need help."	Reworded the question: "Can I figure out what a struggling reader needs to be successful."
10. Create intervention plans to help struggling readers?	"This goes back to the reading specialist. I'm not equipped to make any type of plans for students; that'd be 100+ plans. I don't have time for that."	"I understand the question, but in no way do I see this as part of my job. Students should know how to read but that's what English is for."	"I am not sure if I'd even know how to start this process. There's a toolbox in the front of the book that I have them go to when they're confused."	"This one is tough because our reading specialist is in charge of creating intervention plans. But, I do know how to work with her and how to use them."

12. Determine why a student is having difficult with reading?	“Again, I’m not sure how I would answer this one, it’s not something I really think about. We have the book and we have to use it.”	“I don’t really teach reading in my class; it’s math.”	“I don’t concentrate so much on reading because I know that they do that in English; figure they’ll get it somewhere.”	Reworded the question: “How well can I really troubleshoot why a student is having a hard time reading.”

Note. YRS = Years of Experience. G = Gender. Item numbers are from original STERI for cognitive interviews form.

Difficulty in wording of items. The final category of items were those that participants described as difficult; "unclear"; or "confusing" based on the wording of the items. Overall three items fell into the category of confusing based on wording; two from Form A and one from Form B. Of these three items, two (one from each form) were considered "vague" or confusing by all four participants assessing that form. The third item, from Form A, was considered difficult based on how participants answered the item. Finally, two items had wording slightly altered to alleviate the confusion that one participant had over terminology (see Table 23).

The first type of revision occurred for the items that were described as vague or confusing by participants. On Form A the item, "In your content area, how well can you motivate students to read?" was responded to by participants with questions such as, "Motivate them to read what?"; "For enjoyment, for comprehension?"; and comments such as "This seems vague" or "I'm not sure what you're really asking me with this one" (see Table 23 for complete examples). All four participant responses demonstrated that the item lacked clarity which made it difficult for participants to respond accordingly. Thus, the item was changed to, "In your content area, how well can you motivate students to read *about course topics*?" in an attempt to clarify the type of reading being assessed with the item.

A similar pattern in responses occurred on Form B for the item, "In your content area, how well can you communicate the importance of comprehension?" Three of the four teachers (i.e., social studies, science, and English) responded with questions such as, "Comprehension for what?"; "Comprehension of texts, of reading strategies, of what?";

and comments such as, “I find this question confusing” and “The question is somewhat vague.” The mathematics teacher reworded the item and then stated, “I don’t really deal in terms of reading in my class…” (Teacher E, 20 years of experience, female) Based on the responses of three of the four teachers, it was determined that this item was too vague and therefore difficult to answer. The item was revised, “In your content area, how well can you communicate the importance of *reading* comprehension?” The purpose of adding the term reading before comprehension was to specify the type of comprehension the item was addressing.

The second type of revision occurred for the single item on Form A that participants had trouble answering in its entirety. The item, “In your content area, how well can you modify instruction to motivate students to read?” became problematic as three out of the four participants (i.e., social studies, science, and English) answered only the first part of the item. For example two of teachers stated, “Yes I’m always differentiating my instruction” (Teacher A, science, 11 years of experience, male) and “You have to constantly change up your instruction to keep them [students] engaged” (Teacher D, English, 22 years of experience, female). However, when the follow up probe used was, “are you modifying your instruction to motivate students to read?” participants were unable to address the reading aspect of the item. Instead they responded with examples such as, “Yes, I’m changing up the instruction to motivate students” (Teacher A, science, 11 years of experience, male) or “Yes, like I said if you don’t change up instruction than students are not engaged” (Teacher D, English, 22 years of experience, female). In both examples, the participant is addressing modifying instruction and motivation, but is failing to acknowledge the reading component of the item. Therefore, the item was changed to, "How well can you modify instruction to *encourage* students to read in your content area?" to attempt to alleviate issues surrounding the term motivation.

Two final changes to the items on the STERI were based on Teacher A's (science, 11 years of experience, male) confusion over the terminology of "individual reader" which occurred in one of the Form A items. Although he was the only teacher of the eight that was confused by this terminology (i.e., it appeared in both forms), the meaning of the term was clarified by changing the item to say, "readers of different abilities" instead of "individual readers." Additionally, based on this teacher's confusion, one item on Form B was modified from using the term "individual readers" to state, "individual students." Changing the term *readers* to *students* may make the item more clearly for some teachers, and that this change could be made without losing the aspect of reading from the item itself.

Table 23

Items from STERI Teachers Found Difficult Based on Item Wording

Items	**Content Area of Teacher**			
Form A	Social Studies ***Teacher H*** 3 years of exp. Male	Mathematics ***Teacher F*** 1 year of exp. Female	Science ***Teacher A*** 11 years of exp. Male	English ***Teacher D*** 22 years of exp. Female
2. Modify instruction to motivate students to read?	"Yes, I modify my instruction as much as I can, but it is very time consuming." Follow up: "Yes, but do you modify instruction to motivate students to read?" Teacher: "Well if I'm switching up instruction I should be motivating them to learn."	"Since I don't focus on reading in my classroom I don't change my instruction to deal with reading. In math I do, but not in reading."	"Yes I'm always differentiating my instruction." Follow up: "So you're changing your instruction to encourage them to read?" Teacher: "Yes I'm changing up instruction to motivate them."	"You have to constantly change up your instruction to keep them [students] engaged" Follow up: "Are you modifying your instruction to motivate students to read?" Teacher: "Yes, like I said if you don't change up instruction than students are not engaged."
5. Change your instruction to meet the needs of individual readers?			"I'm not sure what you mean by individual readers; individual levels of readers? If that's what	

			you’re asking then the question is how can I change up my instruction to help individual students”	
11. Motivate students to read?	“Are you talking about academic reading, what type of reading? For enjoyment? For comprehension?”	“I’m not sure what you’re really asking me here.”	“This seems vague. Motivate them to read to get the grade?”	“Motivate them to become lifelong learners? Intrinsically motivate them? Can any teacher really motivate a student to do anything?”
Form B	Social Studies ***Teacher G*** 9 years of exp. Female	Mathematics ***Teacher E*** 20 years of exp. Female	Science ***Teacher B*** 3 years of exp. Female	English ***Teacher C*** 4 years of exp. Female
9. Communicate the importance of comprehension?	“I find this question confusing, Comprehension for what?”	Rewording of item: “Do I tell students how important comprehension is. I don’t really deal in terms of reading in my class, but I tell them math is important.”	“Is this just them knowing it [the material]? The question is somewhat vague.”	“Comprehension of texts, of reading strategies, of what? I’m not sure what you’re asking me here?

Note. YRS = Years of Experience. G = Gender. Item numbers are from original STERI for cognitive interviews form.

Summary of initial cognitive interviews. Based on the eight cognitive interviews, five items (i.e., three on Form A, two on Form B) from the STERI were revised to reflect the teachers' concerns over clarity and wording of items (see Table 24). At the end of the preliminary cognitive interviews, the STERI contained 31 items: nine for the Using Content-Area Reading Strategy Instruction factor, 10 for the Motivating Students to Read factor, and 12 for the Teaching Struggling Readers factor, (see Appendix T). Furthermore, during the cognitive interview process it was noted that some participants were dropping the stem of the items (i.e., In your content area, how well can you…) which changed how they responded to the items on the survey. Therefore, each page of the SurveyMonkey survey (i.e., the survey used for the exploratory analysis) contained the stem of the items at the top and included only five to seven items per page. This was altered primarily to remind teachers of the stem throughout the survey, with the intention that they include it before each item.

Table 24

Items Revised from Forms A and B of STERI Based on Cognitive Interviews

Original Items	**Content Area of Teacher**				**Revised Item**
Form A	Social Studies ***Teacher H*** 3 years of exp. Male	Mathematics ***Teacher F*** 1 year of exp. Female	Science ***Teacher A*** 11 years of exp. Male	English ***Teacher D*** 22 years of exp. Female	
2. Modify instruction to motivate students to read?	"Yes, I modify my instruction as much as I can, but it is very time consuming." Follow up: "Yes, but do you modify instruction to motivate students to read?" Teacher: "Well if I'm switching up instruction I should be motivating them to learn."	"Since I don't focus on reading in my classroom I don't change my instruction to deal with reading. In math I do, but not in reading."	"Yes I'm always differentiating my instruction." Follow up: "So you're changing your instruction to encourage them to read?" Teacher: "Well I differentiate with my reading materials and how I teach to help get the grade."	"You have to constantly change up your instruction to keep them [students] engaged" Follow up: "Are you modifying your instruction to motivate students to read?" Teacher: "Yes, like I said if you don't change up instruction than students are not engaged."	*Modify instruction to encourage students to read?*

5. Change your instruction to meet the needs of individual readers?			“I’m not sure what you mean by individual readers; individual levels of readers? If that’s what you’re asking then the question is how can I change up my instruction to help individual students”		*Change your instruction to meet the needs of readers of different abilities?*
11. Motivate students to read?	“Are you talking about academic reading, what type of reading? For enjoyment? For comprehension?”	“I’m not sure what you’re really asking me here.”	“This seems vague. Motivate them to read to get the grade?”	“Motivate them to become lifelong learners? Intrinsically motivate them? Can any teacher really motivate a student to do anything?”	*Motivate students to read about course topics?*

Form B	Social Studies ***Teacher G*** 9 years of exp. Female	Mathematics ***Teacher E*** 20 years of exp. Female	Science ***Teacher B*** 3 years of exp. Female	English ***Teacher C*** 4 years of exp. Female	**Revised Items**
8. Select course readings to match the needs of individual readers?					*Select course readings to match the needs of individual students?*
9. Communicate the importance of comprehension?	"I find this question confusing, Comprehension for what?"	Rewording of item: "Do I tell students how important comprehension is. I don't really deal in terms of reading in my class, but I tell them math is important."	"Is this just them knowing it [the material]? The question is somewhat vague."	"Comprehension of texts, of reading strategies, of what? I'm not sure what you're asking me here?	*Communicate the important of reading comprehension ?*

Note. YRS = Years of Experience. G = Gender. Item numbers are from original STERI for cognitive interviews form.

Second Round of Cognitive Interviews

After the preliminary cognitive interviews, five of the 31 item STERI were revised (see Table 24). In order to determine if these changes were sufficient in clarifying the meaning of said items, a second round of cognitive interviews occurred. In the second round of cognitive interviews, four teachers (i.e., one from each content area: science, social studies, mathematics, and English) were interviewed using a third form of the STERI for cognitive interviews (Form C; see Appendix K). Teacher and classroom demographic data are located in Tables 8 and 9.

The 12 items included on this form of the STERI were: five items that were changed after the preliminary expert review, three items the previous participants found easy, one item some of the previous participants found difficult due to implementation, and three items that previous participants demonstrated understanding through their responses. These specific items were chosen so that participants were not cognitively drained by only responding to items that were problematic in the preliminary round of cognitive interviews.

The procedure for the second round of cognitive interviews followed the same format as the preliminary cognitive interviews (see Chapter 3 for detailed procedures). However, as these were more focused interviews (i.e., the main focus of these interviews was to determine if the five revised items were more clear to these participants), one anticipated probe (i.e., scripted in advance; Beatty & Willis, 2008; Conrad & Blair, 2004; Willis, 2005) "How could this question be worded differently to make it easier to respond?" was designed to help further clarify any items that were still problematic.

Following a similar pattern as the preliminary cognitive interviews, items were only deemed in need of revision if at least two (i.e., half) of the participants articulated some difficulty with item wording, clarity, or meaning.

Of the 12 items included on the STERI Form C, participants responded to the seven easy items similarly to previous participants. In that, participants demonstrated understanding, articulated items were clear, and reworded items in their own terms. As the main purpose of this second round of cognitive interviews was to examine the revisions made to the five items from the preliminary cognitive interviews; these items will be discussed in detail. Of the five revised items from the preliminary expert review, four of the items were understood by most, if not all of the participants, and one item was still considered problematic by three of the four participants.

Demonstrated understanding of items. Participants demonstrated understanding of items in three ways: rewording of items, explaining what items were asking, and answering items while providing rationale for their decisions (see Table 25). For example, in response to the item "In your content area, how well can you change your instruction to meet the needs of readers of different abilities?" Teacher L stated, "This is a little harder for me, I'd say I do this somewhat. I try to change my instruction often, and I think I do, but it's not all the time because I have so many students that it is very time consuming to try and meet at their needs" (science, three years of experience, male). In this example Teacher L is demonstrating his understanding of the item by answering it using an anchor score from the STERI, and then also explaining why he responded with that score. The understanding is demonstrated through the explanation of the response as he dissects the question and articulates how his answer addresses all aspects of the item. These same types of responses occurred with the social studies, science, and English teachers for the other three items that were previously revised, with none of these teachers reporting any issues (see Table 25). Therefore, four of the five revised STERI items were clear and able to be understood by participants.

Whereas the social studies, science, and English teachers all responded to the items in a very similar way, the mathematics teacher did not follow this same pattern. He demonstrated understanding of all of the items on the STERI, but overall seemed argumentative that any of the items really addressed his classroom context. Further, of all 12 cognitive interview participants (i.e., eight from the preliminary and four from the second round), Teacher I was the most vocal about seeing the items as outside of his job description saying things such as, “I teach math,” “this doesn’t apply to me,” and “that’s not my job” (mathematics, two years of experience, male) throughout the interview. However, even though he expressed his belief that the items were not applicable to a math classroom, he demonstrated meaning through his rewording of items and then subsequent explanation of how the items were asking him about tasks outside of his job description (see Table 25).

Problematic item. One item on the STERI, “In your content area, how well can you motivate students to read about course topics?” was problematic for the social studies, science, and English teachers. In a similar pattern to the preliminary cognitive interviews, participants responded with questions such as, “what type of motivation?”, “read for pleasure?” and were generally unsure of the type of motivation and reading addressed by this item (see Table 25). For example, Teacher J stated, “You’re asking me if I can get students to read about history topics. The answer is yes, but the only motivation I can offer is holding a grade over their head. I’m not sure that’s the type of motivation you’re asking about” (social studies, seven years of experience, female). When the follow up probe, “Well it’s asking, can you get students to read in your class?” was asked, Teacher J responded, “I tell them their grade depends on it, and if that doesn't work they just don't read and generally end up failing.” In this example, Teacher J understands the question, but is not exactly sure if she was addressing the question in the way it was intended.

Whereas the social studies and mathematics teachers were able to answer the item, both the science and English teachers struggled to make meaning of the item overall. After both participants responded that the item was unclear, the anticipated probe, “How could this question be worded differently to make it easier to respond?” was used to determine if participants could suggest a possible revision to aid in item clarity. Teacher L suggested, “Motivate students to participate in reading about course topics” (science, three years of experience, male); however, as participate can have multiple meanings for different people, this suggestion was not used in the revising process. Additionally, Teacher K gave the suggestion of defining the type of motivation as either “intrinsic or extrinsic” (English, nine years of experience, female), however this was a problematic suggestion as some survey participants may not be able to differentiate the two, making the item even less clear. Therefore, although both participants stated the item was still unclear, neither was able to articulate a revision to the item that would increase clarity.

Table 25

Responses from the Second Round of Cognitive Interview Participants on Items Changed after Round One of Interviews

Items	**Content Area of Teacher**			
Items understood	Social Studies ***Teacher J*** 7 years of exp. Female	Mathematics ***Teacher I*** 2 years of exp. Male	Science ***Teacher L*** 3 years of exp. Male	English ***Teacher K*** 9 years of exp. Female
4. Change your instruction to meet the needs of readers of different abilities?	"I'd answer this somewhat to quite a bit. I try and change up how I teach to help all my students and try and meet them where they're at."	Rewording question: "Can I change the way I'm teaching to meet all different ability readers. Well this isn't English class so yeah I change up how I teach math, but it has nothing to do with reading."	"This is a little harder for me, I'd say I do this somewhat. I try to change my instruction often, and I think I do, but it's not all the time because I have so many students that it is very time consuming to try and meet all their needs."	"I feel like this is what it means to teach English. It's all reading, all different types of texts, and if they can't understand what they're reading than they won't do it. So I'm always switching up what and how we read to keep all different levels of readers interested in what we're doing."

6. Select course readings to match the needs of individual students?	"I differentiate a lot of the time because my students have difficulties with the textbook. It's very dense and they find it boring, so I'm constantly bringing in supplemental materials to keep them engaged and help them understand the content."	"I teach math. I use the textbook, that's what we have so no I don't choose different readings outside of the book."	"I'd say this one is quite a bit. I'm always going to the internet to try and find outside material they can use to get the content. Since science is so much about content, they really need to read to understand."	"I'd say this is a little less for me. More so because we have things they have to read, so I can't necessarily change what we read. That's why I change up how we read."
7. Communicate the importance of reading comprehension?	"I'd say I do this quite a bit if not completely because I'm always talking about history being about reading. You can't understand the past if you can't read about it."	"I don't do a ton with reading so this doesn't apply to me, but I understand what the question says…do I tell students reading comprehension is important" Follow up: "Okay, but can you expand on what the item is asking?" Teacher: "It's like I said, do I tell kids that it's important to understand what you read; I don't because	"Science is all about content and they get some of that content from doing, like with experiments and what not, but they have to know how to read and understand the content. So I'm always holding their labs over their heads, like if you can't read about the course topics than you can't do the labs."	"English is all about reading comprehension, from the title to the very last page. I talk about the importance of reading comprehension consistently and believe it's the fundamental principal behind being successful in life. You have to know how to read."

		that’s not my job, but yes I get the question.”		
9. Modify instruction to encourage students to read?	“I try to change how I teach things and the types of activities we do to keep them interested in the topics and peak their interest in reading. I do a lot of pre-reading activities where they predict events etc. to kind of get them hooked before we even get to the reading.”	“Since we don’t read a ton I’m not changing how I teach to encourage them to read. It’s math class, there’s not a ton of reading; if this was English than it’d be a different story.”	“Yes, I am often finding alternative sources on the internet, and having them read newspaper articles and magazines. Our textbook is really difficult so I’m always switching things up to help them get the science content.”	“I think this one goes along with the other question about meeting their needs. As I’m meeting their needs, I’m also encouraging them to read. I can see how the questions are different, but I think I do each one well because I don’t just modify instruction for lower readers, but for my more advanced students as well.”

Item with issues	Social Studies	Mathematics	Science	English
11. Motivate students to read about course topics?	"You're asking me if I can get students to read about history topics. The answer is yes, but the only motivation I can offer is holding a grade over their head. I'm not sure that's the type of motivation you're asking about" Follow up: "Well it's asking, can you get students to read in your class?" Teacher: "I tell them their grade depends on it, and if that doesn't work they just don't read and generally end up failing."	Rewording question: "How well can I encourage students to read about course topics in my class. Again, this is math so they read the textbook and do the problems. There's not a lot of encouraging because you either read it or you don't. No matter what you have to do the problems. "	"I understand the question, but what type of motivation do you mean, for a grade, to actually want to read to learn, the type of motivation you are asking about is not clear."	"Are you asking if they read for pleasure; if they read because they want to; if they read because they have to? I'm not sure what kind of motivation you're talking about."

Note. YRS = Years of Experience. G = Gender. Item numbers are from the STERI for the second round of cognitive interviews form.

Summary of second round of cognitive interviews. After the second round of cognitive interviews, one item "In your content area, how well can you motivate students to read about course topics?" was still reported as problematic. This item was initially revised after the preliminary cognitive interviews to specify to the type of reading (i.e., about course topics) hoping to achieve more clarity. However, the second round of cognitive interviews showed that of the four participants: Teacher I (mathematics, two years of experience, math) was able to answer the item with no issues, Teacher J (social studies, seven years of experience, female) was able to answer the item, but was not sure exactly what the item was asking, and Teachers K (English, nine years of experience, female) and L (science, three years of experience, male) were unable to answer the item and thought it was confusing. Unfortunately none of the participants were able to provide revisions to the items that would enhance clarity. After consulting with the two aforementioned educational psychologists, it was determined that the item would stay in the finalized version of the STERI for exploratory analysis to examine how the item functioned within the factor analysis.

After the second round of cognitive interviews none of the items on the STERI were revised. Therefore, the STERI contained 31 items: nine for the Using Content-Area Reading Strategy Instruction factor, 10 for the Motivating Students to Read factor, and 12 for the Teaching Struggling Readers factor (see Appendix U). Upon completion of all qualitative analyses and revisions to the STERI (i.e., two rounds of expert review and two rounds of cognitive interviews) both face and content validity had been established.

Thus, the 31-item survey was ready to be analyzed using EFA to determine the latent factor structure of the instrument.

Exploratory Factor Analyses

EFA is the most common technique used in scale development and uses variance to determine how items are functioning together within an instrument (Comrey, 1988; DeVellis, 2003). As factor analysis uses measures of variance it is extremely sensitive to outliers. Therefore, before analyses were conducted all data were cleaned to ensure no outliers were present (Dimitrov, 2012). Descriptive statistics were reported for each of the established convergent and discriminant validity scales. Then, the data were subjected to the seven-step EFA process outlined in detail in Chapter 3.

Data cleaning. In order to examine the data for outliers, standardized z-scores were calculated for each scale used (i.e., STERI, Teacher Sense of Efficacy Scale [TSES], Goal Orientations for Teaching [GOT], and Invitations to Self/Others). As z-scores are a measure of standard deviation, any score greater than ± 3.00 could be considered an outlier and should be explored more thoroughly. For the 31 items on the STERI scale, z-scores ranged from -2.84 to 1.99, thus none of the STERI items were considered outliers. The eight items on the TSES subscales had z-scores that ranged from -4.03 to 1.8, with 14 z-scores that could be considered outliers, ranging from -3.02 to -4.34. All 14 z-scores outside of the ± 3.00 range were visually inspected and histograms created. I found no reason to delete any of these scores as they were reflections of participants' responses and none were true outliers (i.e., no participant chose the same response category for the entire survey). For the eight items on the GOT subscales, z-scores ranged from -.4.11 to 1.52, with six z-scores outside of the ± 3.00 that ranged from -3.20 to -4.11. After inspection, none of these scores were deleted because they reflected participants' valid responses. Finally, for the 12 items on the Invitations to Self/Others scale z-scores ranged from -3.72 to 1.58, with 15 scores outside of the ± 3.00 range. These scores ranged from -3.06 to -3.72, however upon review no scores were deleted as they represented participants' actual responses. Thus, after examining z-scores, the data was free from outliers and further analyses could be conducted.

Descriptive statistics. Mean composites were created for all established subscales (i.e., TSES, GOT, and Invitations scales). Composites were created by taking the average from multiple indicators into a single score that represents the underlying construct (DeVellis, 2013). To create mean composites, I summed the item scores for each item on a particular subscale and then divided that by the total number of items on the scale [e.g., Instructional Strategies: (item 5) + (item 9) + (item 10) + (item 12) = / 4]. The purpose of creating these composites was to compare scores on subscales, instead of individual items (i.e., each subscale has one score that is averaged from all item scores comprising that subscale). Two mean composites were created for the TSES subscales, two for the GOT subscales, and two for the Invitations subscales.

***Convergent scales*.** For the TSES, items were measured on a nine-point scale. Two mean composites were created for each of the two TSES subscales: Engagement and Instructional Strategies. The mean of the four items on the TSES Engagement factor was 6.35, with a standard deviation of 1.30, and a reliability of $\alpha = .81$. For the four items on the TSES Instructional Strategies factor, the mean was 6.89, with a standard deviation of

1.34, and a reliability of $\alpha = .86$. These two subscales were moderately positively correlated, $r = .51$, $p < .01$.

For the GOT, items were measured on a five-point scale. For the GOT, two mean composites were created for each of the two subscales: Mastery and Relational Goals. For the four items on Mastery Goals factor, the mean was 4.18, with a standard deviation of .64, and a reliability $\alpha = .81$. The four items on the Relational Goals factor had a mean of 3.60, a standard deviation of .94, and a reliability of $\alpha = .93$. These subscales were weakly positively correlated $r = .33$, $p < .01$.

***Discriminant scales*.** For the Invitations to Self and Others scales, items were measured on a nine-point scale. Two mean composites were created for each of the two subscales: Invitations to Self and to Others. The six items on the Invitations to Self factor had a mean of 6.42, a standard deviation of 1.45, and a reliability of $\alpha = .87$. The six items on the Invitations to Others factor had a mean of 7.10, a standard deviation of 1.19, and a reliability of $\alpha = .85$. Further, the two subscales were weakly positively correlated ($r = .18$, $p < .01$).

EFA. The original 31 items of the STERI were analyzed using EFA following the seven-step process outlined in detail in Chapter 3. In step one, data were cleaned to ensure no univariate or multivariate outliers were present. Step two, Bartlett's Test of Sphericity and Kaiser-Meyer Olkin (KMO) test statistics were examined to ensure the data were acceptable for use with an EFA. Step three, correlational analyses were conducted to determine which extraction method of EFA to use (i.e., Principal Components Analysis [PCA] or Principal Axis Factoring [PAF]). Step four, the method of rotation was decided upon based on the correlational analysis conducted in step three. Step five, examination of eigenvalues and the Scree Plot was used to decide the number of factors to extract. Step six, factor loadings were used in determining deletion and retention of items. Step seven, factor structure was used to interpret factors.

Step one. To ensure the data were appropriate for statistical analyses: means, standard deviations, skewness, and kurtosis were calculated for each item on the STERI. The 31 STERI items were measured on a six-point scale with item means between 3.06 and 4.67 and standard deviations ranging from 1.12 to 1.56. Additionally, skewness and kurtosis analyses revealed no extreme values (i.e., |3| for skewness and |10| for kurtosis, see Table 26; Kline, 1998).

Step two. To determine the suitability of the data for an EFA, Bartlett's test of Sphericity and the KMO measure of sampling adequacy (Kaiser, 1974) were analyzed. Bartlett's test was significant, χ^2 (465) = 1022.26, $p < .001$, and the KMO statistic was .97, well above the recommended value of .6 and suggesting a "marvelous" sampling adequacy (Kaiser, 1974). Additionally, the diagonals of the anti-image correlation matrix were all over .50, supporting the inclusion of each item into the factor analysis. Correlations among items ranged between $r = .41$ and $r = .85$ (ps < .01), signifying moderate to strong positive relationships between items on the STERI.

Table 26

Skewness, Kurtosis, Mean, and Standard Deviation for the Original 31 Item STERI

Item: In your content area how well can you…	Skewness	Kurtosis	Mean	SD
Q3: Help students monitor their use of reading strategies?	-.05	-.99	3.56	1.37
Q16: Teach different reading strategies?	-.16	-1.12	3.62	1.56
Q4: Encourage students to find the reading strategy that works best for them?	-.17	-.99	3.73	1.38
Q18: Help students apply reading strategies to improve comprehension?	-.21	-1.03	3.68	1.42
Q1: Model effective reading strategy use?	-.41	-.84	4.06	1.43
Q7: Assess students' use of reading strategies?	-.11	-.99	3.52	1.42
Q22: Help students choose which reading strategy to use?	-.08	-1.08	3.31	1.35
Q29: Differentiate reading strategy instruction?	-.03	-1.25	3.32	1.49
Q6: Model effective reading for struggling readers?	-.34	-.81	3.98	1.38
Q28: Identify improper reading strategy use?	.13	-1.06	3.06	1.39
Q15: Help struggling readers comprehend text?	-.61	-.23	4.16	1.16
Q14: Inspire students of all reading abilities to read?	-.35	-.42	3.97	1.22
Q20: Encourage uninterested students to read?	-.31	-.53	3.81	1.28
Q17: Emphasize the importance of reading?	-.82	.09	4.67	1.19
Q8: Motivate students to improve their understanding when reading?	-.58	-.00	4.20	1.12
Q23: Increase students' confidence in reading?	-.43	-.72	3.92	1.31
Q19: Use various types of assessment to motivate students to read?	-.41	-.81	3.98	1.41
Q24: Use a variety of texts to encourage students to read?	-.57	-.70	4.11	1.48
Q13: Encourage students to self-assess their reading comprehension?	-.34	-.73	3.73	1.33
Q11: Motivate students to read about course topics?	-.61	-.22	4.13	1.24
Q26: Communicate the importance of reading comprehension?	-.77	.00	4.45	1.25
Q2: Modify instruction to encourage students to read?	-.62	-.26	4.29	1.19
Q31: Help struggling readers improve their comprehension?	-.35	-.71	3.70	1.30

Q27: Adapt reading materials for struggling readers?	-.32	-.76	3.68	1.40
Q30: Help struggling readers self-assess their own comprehension?	-.20	-1.00	3.41	1.35
Q10: Create intervention plans to help struggling readers?	-.06	-.98	3.39	1.41
Q9: Determine the reading needs of struggling readers?	-.38	-.78	3.72	1.33
Q21: Use a variety of assessments to help struggling readers?	-.41	-.71	3.80	1.38
Q12: Determine why a student is having difficult with reading?	-.10	-1.08	3.40	1.41
Q25: Select course readings to match the needs of individual students?	-.31	-1.01	3.70	1.49
Q5: Change your instruction to meet the needs of readers of different abilities?	-.58	-.28	4.19	1.25

Note. SD = Standard Deviation.

Step Three. After it was determined that the data were sufficient for an EFA, the items were included in two different factor analyses: PCA with Varimax rotation and PAF with Promax rotation. The factor correlation matrix revealed factors were positively moderately correlated (Factors 1 and 2: $r = .71$, Factors 2 and 3: $r = .77$, Factors 1 and 3: $r = .77$, ps $< .01$). Based upon these moderate correlations, PAF was chosen as the method of extraction over PCA as PAF allows correlations to be estimated, whereas PCA fixes correlations at 0 (Reise, Waller, & Comrey, 2000). Additionally, PCA assumes the variables account for 100% of the variance, therefore fixing items' initial commonalities at 1; however, PAF does not assume 100% of the variance will be explained, and provides estimates of item commonalities based on the item variability in common with other items. Further, as the goal of the EFA was to determine the factor structure of the STERI, PAF represented a better analysis as it is used to detect the underlying latent structure of an instrument; whereas PCA is used for data reduction (Osborne & Costello, 2005; Reise et al., 2000).

Step four. When determining the type of rotation to use, it was important to again examine the correlations between the hypothesized factors. Varimax rotation assumes orthogonal, or non-correlated factors, whereas Promax rotation allows for the factors to be oblique, or correlated (Comrey, 1988; Osborne & Costello, 2005; Reise et al., 2000). Because the factors were moderately positively correlated (Factors 1 and 2: $r = .71$, Factors 2 and 3: $r = .77$, Factors 1 and 3: $r = .77$, ps $< .01$), an oblique rotation method must be used when examining factor structure. Both Direct Oblimin and Promax rotation methods were initially conducted; however, Promax was decided upon due to the difficulty in interpreting the negative factor loadings in the Oblimin output. As all three factors were positively moderately correlated, my initial hypothesis from Chapter 3 (i.e., all factors would be weakly to moderately positively correlated) was partially confirmed. However, the correlations were stronger than expected.

Step five. Once extraction and rotation method were determined, eigenvalues determined the number of factors to extract. The initial factor analysis saw three factors all with eigenvalues over 1 (Factor 1: 20.64, Factor 2: 1.90, Factor 3: 1.07). Examination of Cattell's scree plot confirmed that three factors were above the "elbow" of the plot. Further, after rotation these three factors accounted for 73.75% of the variance, with Factor 1 explaining 16.69%, Factor 2 accounting for 16.58%, and Factor 3 explaining 17.79% of the variance respectively. Next, item communalities were inspected to ensure that each item had at least half of its variance (.50) accounted for by the factor. All commonalities were between .57 and .84, therefore more than half the variance in each item was accounted for by its extracted factor (see Table 26).

Step six. After examining commonalities, the structure matrix was analyzed to determine the factor loadings of items and to determine if the 3-factor structure was theoretically sound. All factor loadings ranged between .55 - .91, well above the recommended loading of at least .30 (DeVellis, 2003); therefore each item was loading high enough onto at least one factor. Further, items were considered to cross-load if their primary loading was .50 or above and their secondary loading was .50 or higher and within .05 of the primary loading (e.g., factor loadings: .81, .64, .84; this is cross-loading on Factors 1 and 3 as both are above .50 and are within .03 from one another). Upon inspection, items loaded as follow: 6 items loaded onto Factor 1(Items: 3, 16, 4, 18, 1, 7) with factor loadings of .86 to .91, 9 items loaded onto Factor 2 (Items: 14, 20, 17, 8, 24, 13, 11, 26, 2) with factor loadings of .80 to .87, and 7 items loaded onto Factor 3 (Items: 27, 31, 30, 10, 9, 21, 5) with factor loadings of .73 to .89 (see Table 27). The other nine items (Items: 22, 29, 28, 15, 6, 23, 19, 25, 12) either cross-loaded or surprisingly loaded onto a different primary factor than anticipated.

Step seven. After reviewing the items that loaded onto each factor, it was determined that Factor 1 could be interpreted as self-efficacy for Using Content-Area Reading Strategy Instruction, Factor 2 was self-efficacy for Motivating Students to Read, and Factor 3 was self-efficacy for Teaching Struggling Readers. Therefore, the hypothesized three factor structure was found in the initial EFA solution; however this solution was not ideal as eight items were not distinguishing onto one primary factor. Within this solution, five items cross-loaded onto Factors 1 and 3 (Items: 12, 15, 22, 28, 29), three items cross-loaded onto Factors 2 and 3 (Items: 19, 23, 25), and one item loaded unexpectedly on Factor 3 instead of the a priori assumption of Factor 1 (Q6; see Table 27).

Table 27

Factor Loadings for the Original 31 Item STERI

Item: In your content area how well can you…	Comm.	Factor 1	Factor 2	Factor 3
Q3: Help students monitor their use of reading strategies?	.83	**.91**	.66	.68
Q16: Teach different reading strategies?	.84	**.90**	.65	.79
Q4: Encourage students to find the reading strategy that works best for them?	.81	**.90**	.62	.64
Q18: Help students apply reading strategies to improve comprehension?	.85	**.89**	.66	.77
Q1: Model effective reading strategy use?	.80	**.89**	.55	.59
Q7: Assess students’ use of reading strategies?	.76	**.86**	.65	.73
Q22: Help students choose which reading strategy to use?	.80	**.86**[a]	.65	**.81**[a]
Q29: Differentiate reading strategy instruction?	.82	**.85**[a]	.67	**.85**[a]
Q6: Model effective reading for struggling readers?	.70	**.80**[b]	.70	**.74**[b]
Q28: Identify improper reading strategy use?	.70	**.79**[a]	.61	**.78**[a]
Q15: Help struggling readers comprehend text?	.61	**.73** [a]	.65	**.72** [a]
Q14: Inspire students of all reading abilities to read?	.76	.57	**.87**	.64
Q20: Encourage uninterested students to read?	.74	.61	**.86**	.68
Q17: Emphasize the importance of reading?	.73	.70	**.84**	.68
Q8: Motivate students to improve their understanding when reading?	.71	.64	**.84**	.69
Q23: Increase students’ confidence in reading?	.77	.63	**.83**[c]	**.81**[c]
Q19: Use various types of assessment to motivate students to read?	.75	.63	**.83**[c]	**.80**[c]
Q24: Use a variety of texts to encourage students to read?	.73	.60	**.83**	.77
Q13: Encourage students to self-assess their reading comprehension?	.70	.63	**.82**	.74
Q11: Motivate students to read about course topics?	.65	.54	**.80**	.62
Q26: Communicate the importance of reading comprehension?	.69	.69	**.80**	.73
Q2: Modify instruction to encourage students to read?	.68	.70	**.80**	.65
Q31: Help struggling readers improve their comprehension?	.83	.75	.78	**.89**

Q27: Adapt reading materials for struggling readers?	.79	.65	.69	**.89**
Q30: Help struggling readers self-assess their own comprehension?	.80	.68	.75	**.89**
Q10: Create intervention plans to help struggling readers?	.75	.72	.69	**.86**
Q9: Determine the reading needs of struggling readers?	.75	.75	.69	**.85**
Q21: Use a variety of assessments to help struggling readers?	.71	.71	.72	**.83**
Q12: Determine why a student is having difficult with reading?	.66	**.74**[a]	.65	**.77**[a]
Q25: Select course readings to match the needs of individual students?	.63	.55	**.71**[c]	**.76**[c]
Q5: Change your instruction to meet the needs of readers of different abilities?	.57	.65	.64	**.73**

Note. Comm. = Commonalities. Bold = Factor that item loaded onto. [a] = Items that cross-loaded onto Factor 1 and Factor 3. [b] = Item that unexpectedly loaded on Factor 1, instead of the intended Factor 3. [c] = Items that cross-loaded onto Factor 2 and Factor 3.

Although a three-factor structure was found with the initial examination of the data, issues of cross-loading still plagued the initial 31 items. In order to alleviate these issues, I examined why certain items were cross-loading and looked at wording issues which may have caused the difficulties. Distinguishable patterns were found in the wording of items and their subsequent cross-loadings. For the five items that cross-loaded onto both Factor 1 (Using Content-Area Reading Strategy Instruction) and Factor 3 (Teaching Struggling Readers), three items implied struggling readers without referencing them directly. More specifically, Q22: Help students choose which reading strategy to use, Q29: Differentiate reading strategy instruction and Q28: Identify improper reading strategy use, cross-loaded similarly in that all three items were meant to factor onto the Using Content-Area Reading Strategy Instruction factor, however they implied helping struggling readers. Further, teachers could have responded that these items functioned on both factors because helping students choose which strategy to use, differentiating, and using reading strategies improperly may occur most often when students are struggling. Therefore, the wording of these items may have caused them to cross-load onto Factors 1 and 3.

The other two items that cross-loaded onto Factors 1 and 3 (Q15 and Q12) both unintentionally implied reading strategy use. Q15: Help struggling readers comprehend text and Q12: Determine why a student is having difficult with reading, were both written to load onto the Teaching Struggling Readers factor, but both cross-loaded with the Using Content-Area Reading Strategy Instruction factor. This cross-loading may be explained by the underlying assumption of reading strategy use when helping struggling readers.

Teachers may have viewed Q15 as reading strategy instruction and teaching struggling readers because they use reading strategies to help struggling readers. Moreover, Q12 may have been viewed similarly with teachers using reading strategy instruction to help determine the needs of a struggling reader. These items both functioned similarly in that the wording of the items may have caused teachers to see a link between reading strategy instruction and teaching struggling readers.

Another pattern among the cross-loading items was that both Q23 and Q25 cross-loaded on Factor 2 (Motivating Students to Read) and Factor 3 (Teaching Struggling Readers). Item 23: Increase students confidence in reading, was intended to load onto Factor 2; however, its cross-loading onto Factor 3 is understandable because struggling readers are generally the students whose confidence needs to be increased in reading. In addition, item 25: Select course readings to match the needs of individual students, was meant to load onto Factor 3, but its implication for motivating students may have caused participants to respond similarly to items that loaded onto Factor 2. In particular, the differentiation of course readings is a motivating practice and therefore the wording of the item was problematic because teachers responded to this item in a similar way to items that loaded onto Factor 2.

The final problematic item, Q6: Model effective reading for struggling readers unexpectedly loaded onto Factor 1 (Using Content-Area Reading Strategy Instruction) instead of Factor 3 (Teaching Struggling Readers). Even though this item contained the term struggling readers, participants responded to the item similarly to reading strategy items. This surprising loading was most likely because modeling is a type of strategy instruction, therefore the item had poor discrimination between Factors 1 and 3, causing the loading.

Although some of the nine items that cross-loaded had wording issues, it was decided to begin deleting items to determine if a better model fit could be achieved. Further, having 31 items on the STERI for the EFA was intentional for exploratory purposes; however the final version of the scale was never intended to contain 31 items. Having more items than necessary on a scale provides the opportunity to gain the best factor solution possible by deleting items when necessary (Gehlbach & Brinkworth, 2011). EFA is based upon variance, therefore items were dropped from the instrument one at a time as the deletion of one item changed the commonalities and factor loadings of all other items. Therefore, after a deletion of one item, the EFA was conducted again (with PAF extraction and Promax rotation), new results interpreted, and final decisions made about item inclusion and exclusion. Therefore, in order to create a more parsimonious scale, problematic items were deleted from the STERI in lieu of revisions.

As no skewness or kurtosis issues arose when initial examining the data, deletion of items began by choosing the "problem item" with the lowest commonality (i.e., which item of the eight that had the lowest commonality), deleting this item, conducting the EFA, and interpreting the results. This process continued until all combinations of deleted problem items occurred. For example, I started by individually deleting each problem item and then conducting the EFA (i.e., 8 problem items, 8 EFAs), then I deleted two-item combinations (e.g., Q8 and Q10; Q8 and Q11), next I deleted three-item combinations, and so on until I had utilized all forms of item combinations. Even though

I had eight problem items, in finality I deleted nine items to create the best EFA fit statistics and the most psychometrically sound instrument. Thus, after various iterations of EFAs, nine items were deleted (Q29, Q6, Q28, Q15, Q23, Q19, Q24, Q26, Q12). These conclusions were based on the notion that deleting just one item changes the factor structure, commonalities, and factor loadings of all other items

Final solution. After deleting nine items from the STERI, the remaining 22 items were analyzed using EFA. Based on the factors being correlated, PAF was used with a Promax rotation. A Promax rotation was still considered ideal as the three factors were moderately positively correlated (Factors 1 and 2: $r = .68$, Factors 1 and 3: $r = .77$, and Factors 2 and 3: $r = .77$, ps < .01). Bartlett's Test of Sphericity was significant, χ^2 (231) = 7173.29, $p < .001$, and the KMO measure of sampling adequacy was still considered "marvelous" at .97. The diagonals of the anti-image correlation matrix were all over .50, supporting the inclusion of all items into the EFA. Commonalities for each item were between .59 - .84 (see Table 28); thus more than half the variance in each item was accounted for by the factors. Examination of Cattell's scree plot and initial eigenvalues (Factor 1: 14.66, Factor 2: 1.56, Factor 3: 1.00) suggested a three factor solution. Further, the three factor solution accounted for 74.98% of the variance, with almost equal variance explained by each factor (Factor 1: 28.26%, Factor 2: 25.69%, Factor 3: 24.32%).

Additionally, factor loadings were used to determine what items loaded onto which factors. After inspection, the hypothesized three-factor structure was upheld with Factor 1 interpreted as self-efficacy for Using Content-Area Reading Strategy Instruction, Factor 2 as Motivating Students to Read, and Factor 3 as Teaching Struggling Readers. Moreover, the number of items was nearly equal across factors: Factor 1 had seven items (Q3, Q16, Q4, Q18, Q1, Q7, Q22) with factor loadings of .87 to .91 that all pertained to teachers use of reading strategy instruction, Factor 2 had seven items (Q14, Q20, Q8, Q11, Q17, Q13, Q2) with factor loadings of .79 to .89 that dealt with increasing students' confidence and motivating them to read, and Factor 3 had eight items (Q27, Q31, Q30, Q10, Q9, Q21, Q25, Q5) with factor loadings of .76 to .91 that related to teaching and helping struggling readers (see Table 28). More importantly, the 22 item STERI had no cross-loading or unexpected loading items; therefore each item loaded onto only one primary factor and did not have a secondary loading within .05 of any other factor (See Table 28). Thus, the final STERI contained 22 items across the three factors of self-efficacy for: Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers.

Table 28

Factor Loadings for Final 22 Item STERI

Item: In your content area how well can you…	Comm.	Factor 1	Factor 2	Factor 3
Q3: Help students monitor their use of reading strategies?	.83	**.91**	.67	.69
Q16: Teach different reading strategies?	.84	**.90**	.64	.79
Q4: Encourage students to find the reading strategy that works best for them?	.81	**.90**	.62	.65
Q18: Help students apply reading strategies to improve comprehension?	.82	**.90**	.65	.78
Q1: Model effective reading strategy use?	.79	**.88**	.55	.61
Q7: Assess students' use of reading strategies?	.77	**.87**	.65	.73
Q22: Help students choose which reading strategy to use?	.81	**.87**	.64	.80
Q14: Inspire students of all reading abilities to read?	.79	.57	**.89**	.66
Q20: Encourage uninterested students to read?	.77	.59	**.88**	.68
Q8: Motivate students to improve their understanding when reading?	.68	.64	**.82**	.68
Q11: Motivate students to read about course topics?	.67	.53	**.82**	.62
Q17: Emphasize the importance of reading?	.69	.71	**.80**	.70
Q13: Encourage students to self-assess their reading comprehension?	.69	.66	**.80**	.74
Q2: Modify instruction to encourage students to read?	.67	.70	**.79**	.68
Q27: Adapt reading materials for struggling readers?	.83	.65	.66	**.91**
Q31: Help struggling readers improve their comprehension?	.84	.75	.77	**.90**
Q30: Help struggling readers self-assess their own comprehension?	.77	.69	.73	**.88**
Q10: Create intervention plans to help struggling readers?	.75	.73	.68	**.86**
Q9: Determine the reading needs of struggling readers?	.75	.75	.67	**.85**
Q21: Use a variety of assessments to help struggling readers?	.70	.71	.70	**.83**
Q25: Select course readings to match the needs of individual students?	.62	.55	.68	**.77**
Q5: Change your instruction to meet the needs of readers of different abilities?	.59	.65	.64	**.76**
Eigenvalues		14.66	1.56	1.00
% of variance explained		28.26	25.69	24.32
Cronbach's alpha		.96	.94	.95

Note. Comm. = Commonalities. Bold = Highest factor loading.

Descriptive Statistics of the STERI

To examine if contextual factors had any relation to teachers' reported self-efficacy on the overall STERI as hypothesized in Chapter 3, one-way Analysis of Variance (ANOVAs), independent *t*-tests, and chi-square test of independence analyses were conducted. As the survey instrument was designed with drop down menus to alleviate participants incorrectly entering data, for many of the analyses groups needed to be created so that responses could be compared. Therefore, whenever too many individual categories were present, I combined similar responses into groups in order to conduct different analyses. Both teacher and environmental factors were examined to determine if they contributed to teachers' responses on the STERI.

Teacher factors. The teacher factors that were examined were: gender, years of experience, highest level of education, preparation (i.e., number of literacy course taken for credit and number of professional development courses), and ethnicity/race. One-way ANOVAs, independent *t*- tests, and chi-square analyses were used to determine if these factors related to teachers' responses on the STERI.

Non-significant factors. Two factors did not have any significant differences: years of teaching experience and race/ethnicity. In order to examine years of experience, responses were grouped as: 1-5, 6-10, 11-15, 16-20, and 20+ in order to collapse individual responses into more equal sample sizes. A one-way ANOVA showed that years of experience was non-significant, $F(4, 257) = .17$, $p > .05$. When examining the effect of ethnicity on teachers' performance on the STERI, the sample was grouped as "white" and "non-white" in order to make the groups comparable based on sample size.

An independent t-test showed that race was also non-significant, $t(283) = 1.09$, $p > .05$. Therefore, both years of experience and race/ethnicity did not contribute to teachers' responses on the STERI, aligning with findings from similar teacher self-efficacy studies (e.g., Raudenbush et al., 1992; Tschannen-Moran & Johnson, 2011).

Significant factors. Three teacher factors were significant when examining their relation to teachers' responses to the STERI: gender, highest level of education, and preparation. Although gender, highest level of education, and preparation have not consistently been reported as contributing to participants' responses to self-efficacy studies, all three factors related to participants' responses on the STERI.

Gender. One teacher factor that demonstrated difference in scores on the STERI across groups was gender. This finding is concurrent with other studies (e.g., Klassen & Chiu, 2010; Raudenbush et al., 1992; Ross et al., 1996) that found gender had a significant effect on participants' self-efficacy. In this analysis, Levene's Homogeneity of Variance was not violated, however it is important to note that there were 184 females as opposed to 100 males included in the analysis. An independent *t*-test showed that females reported a statistically significantly higher mean ($M_{\text{diff}} = .56$) on the STERI than males, $t(282) = 4.26$, $p < .001$, $d = .54$. A chi-square test of independence was conducted on teachers' primary teaching assignment and their gender. This analysis revealed there was a relationship between content area taught and gender, χ^2 (3, $N = 317$) = 60.13, $p < .001$. This was most notably seen when examining the gender of English and social studies teachers. The expected count of English teachers was 79.1 female and 40.9 male; however the actual count was 106 female and 14 male. Further, the expected count of social studies teachers was 40.9 female and 21.1 male; however the actual count was 20 female and 42 male. Therefore, this sample had more female English and male social studies teachers than what was to be expected; thus, gender being a significant variable may have been dependent on the sample used in this study.

Level of education. Another teacher factor that showed significant difference across groups on the STERI was teachers' level of education; a contradictory finding to the work of Raudenbush and colleagues (1992) and Ross and colleagues (1996) who found level of education did not contribute to teacher self-efficacy. Initially, a one-way ANOVA was performed on the groups as "bachelors and bachelors plus," "masters and masters plus," and "doctorate"; however, upon examination of the results the doctorate group had only 15 participants, whereas the bachelors group had 60, and the masters group had 206. Therefore, it was determined that the ANOVA results could be skewed based on the vast difference in sample size; thus, the groups were combined to make more equally sized groups. In order to more accurately compare groups based on sample size, categories were condensed into "bachelors and bachelors plus" and "masters, masters plus, and doctorate." An independent *t*-test was conducted, Levene's Homogeneity of Variance was upheld ($p > .05$), and mean difference in level of education was found to be significant, $t(284) = -2.20$, $p < .05$, $d = .31$, with the masters plus group scoring higher on the STERI than the bachelors plus group ($M_{\text{diff}} = .34$). Thus, teachers' with a higher level of education (i.e., masters, masters plus, and doctorate) scored statistically significantly higher on the STERI than did teachers with less education (i.e., bachelors, bachelors plus).

Preparation. A final significant teacher factor that contributed to participants' responses on the STERI was their preparedness for teaching literacy (i.e., number of literacy courses taken for credit and number of literacy professional development sessions attended). These findings align with Tschannen-Moran and Johnson who found that professional development experiences contribute to teachers' self-efficacy (2011). For both preparedness variables, as teachers were able to choose from a drop-down menu and scores varied from 0-31 courses, five groups were created in order to conduct analyses (i.e., 0, 1-5, 6-11, 12-20, 21+). Two separate one-way ANOVAs were run to assess each of the preparedness variables contribution to teachers' self-efficacy as reported in the STERI.

For the first analysis, relationship of number of literacy courses taken for credit, Levene's Homogeneity of Variance was upheld and statistically significant differences were reported, $F(4, 275) = 15.75$, $p < .01$, $\eta^2 = 22.90\%$. Therefore, the number of literacy courses taken for credit related to participants' scores on the STERI and a Tukey post-hoc test was conducted to determine which group differences were statistically significant. Teachers with 21+ literacy courses had a statistically significantly higher mean score on the STERI than teachers with no literacy courses ($M_{\text{diff}} = 1.41$, $p < .01$, $d = 1.55$). Those teachers with 12-20 courses had a statistically significantly higher mean than both teachers with 0 courses ($M_{\text{diff}} = 1.49$, $p < .01$, $d = 1.54$) and teachers will 1-5 courses ($M_{\text{diff}} = .98$, $p < .01$, $d = 1.03$). Additionally, teachers with 6-11 courses had a statistically significantly higher mean than teachers with 0 courses ($M_{\text{diff}} = 1.48$, $p < .01$, $d = 1.64$) and teachers with 1-5 courses ($M_{\text{diff}} = .97$, $p < .05$, $d = 1.10$). Further, teachers with 1-5 courses had a statistically significantly higher mean than those teachers with 0 courses ($M_{\text{diff}} = .50$, $p < .05$, $d = .49$). Therefore, the number of literacy courses taken for credit contributed teachers' scores on the STERI; teachers that took courses for credit had a higher mean score on the STERI than teachers with no courses. See Table 29 for complete information across STERI subscales.

In the second analysis, number of literacy professional development sessions attended, Levene's Homogeneity of Variance was violated ($p < .05$) and statistically significant differences were found, $F(4, 281) = 15.93$, $p < .01$, $\eta^2 = 22.67\%$. Thus, the number of literacy professional development courses participants attended related to their scores on the STERI (see Table 30). As Levene's Homogeneity of Variance was violated, a Games-Howell post-hoc analysis was conducted to determine where group differences occurred. Teachers with 21+ literacy professional development sessions had statistically significantly higher mean scores on the STERI than teachers with 0 sessions ($M_{\text{diff}} = 1.72$, $p < .01$, $d = 2.07$), teachers with 1-5 sessions ($M_{\text{diff}} = 1.12$, $p < .01$, $d = 1.35$), and teachers with 6-11 sessions ($M_{\text{diff}} = .62$, $p < .05$, $d = .78$). Those teachers with 12-20 literacy professional development sessions had a statistically significantly higher mean score than teachers with 0 sessions ($M_{\text{diff}} = 1.31$, $p < .01$, $d = 1.31$) and those with 1-5 sessions ($M_{\text{diff}} = .71$, $p < .05$, $d = .71$). Further, teachers with 6-11 sessions had a statistically significantly higher mean than teachers with 0 sessions ($M_{\text{diff}} = 1.10$, $p < .01$, $d = 1.10$) and teachers with 1-5 sessions ($M_{\text{diff}} = .50$, $p < .05$, $d = .50$). Finally, teachers with 1-5 sessions had a statistically significantly higher mean score than teachers with 0 sessions ($M_{\text{diff}} = .60$, $p < .01$ $d = .58$). Hence, teachers who had attended literacy professional development sessions had higher mean scores on the STERI than those teachers who had not attended any sessions.

Table 29

Group Mean Differences Based on Number of Literacy Courses

	Overall STERI	Using Reading Strategies	Motivating Student to Read	Teaching Struggling Readers
Number of Literacy Courses	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
0	3.25 (1.03)[a]	2.95 (1.11)[a]	3.69 (1.12)[a]	3.08 (1.56)[a]
1-5	3.75 (1.01)[b]	3.56 (1.23)[b]	4.11 (0.98)[ab]	3.68 (1.10)[b]
6-11	4.72 (0.74)[c]	4.78 (0.84)[c]	4.74 (0.73)[bc]	4.59 (0.83)[c]
12-20	4.73 (0.89)[c]	4.88 (0.97)[c]	4.68 (0.89)[b]	4.65 (0.91)[c]
20+	4.66 (0.76)[c]	4.62 (0.90)[bc]	4.89 (0.58)[c]	4.50 (0.93)[b]

Note. Means sharing a letter in their superscript are not significantly different at the .05 level according to a Tukey HSD post-hoc test.

Table 30

Group Mean Differences Based on Number of Literacy PD Courses

	Overall STERI	Using Reading Strategies	Motivating Students to Read	Teaching Struggling Readers
Number of PD Courses	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
0	3.10 (1.03)[a]	2.80 (1.19)[a]	3.67 (1.14)[a]	2.97 (1.07)[a]
1-5	3.70 (1.03)[b]	3.51 (1.22)[b]	4.00 (1.02)[ab]	3.61 (1.23)[b]
6-11	4.20 (0.97)[c]	4.13 (1.84)[c]	4.44 (0.92)[c]	4.12 (1.05)[c]
12-20	4.41 (0.97)[cd]	4.50 (0.93)[c]	4.55 (0.87)[cd]	4.25 (1.18)[bc]
20+	4.82 (0.57)[d]	4.78 (0.69)[c]	4.90 (0.44)[d]	4.69 (0.81)[c]

Note. Means sharing a letter in their superscript are not significantly different at the .05 level according to a Games-Howell post-hoc test.

Environmental factors. The environmental factors examined were: number of courses taught, grade level taught, Title I school status, percent of students that were English Language Learners (ELLs), percent of students with Individualized Education Plans (IEPs), students' reading level, and teachers' primary content area of instruction. For all variables quantified by a percentage, the groups were predetermined on the survey as: 0, 1-25, 26-50, 51-75, and 76-100. In order to examine differences, one-way ANOVAs, intendent *t* - tests, and chi-square test of independence analyses were conducted.

Non-significant factors. When examining environmental factors, the number of courses teachers taught was non-significant, $F(7, 275) = 1.70$, $p > .05$. In addition, the percent of students that were ELLs was non-significant, $F(4, 278) = .17$, $p > .05$, as well as the percent of students with IEPs, $F(4, 277) = 1.58$, $p > .05$. Further, the percent of above-grade was not significant, $F(4, 278) = 1.07$, $p > .05$, nor were the percent of on-graders readers, $F(4, 276) = 2.07$, $p > .05$. Therefore, the number of courses a teacher taught, the percent of students whom were ELLs, the percent of students with IEPs, and the percent of above-grade and on-grade readers that a teacher instructs does not have a statistically significant relationship to teachers' responses on the STERI.

Significant factors. Factors that had a statistically significant contribution to teachers' self-efficacy as reported on the STERI were: grade level taught, school Title I status, percent of below-grade readers, and teachers' primary content area of instruction. To examine differences, one-way ANOVAs, intendent t – tests, and chi-square test of independence were conducted.

Grade level. In an effort to make grade level taught comparable across the sample, two groups were created: "single grade level" and "multiple grade levels" based on the number of grade levels a teacher taught. Levene's Homogeneity of Variance was upheld, and an independent t-test revealed that there was a significant mean difference, $t(279) = 3.27$, $p < .01$, $d = .46$, between teachers teaching one grade level and those teaching multiple grade levels. This analysis revealed that teachers teaching a single grade level had a higher mean score on the STERI ($M_{\text{diff}} = .49$) than teachers teaching multiple grade levels. Further, a chi-square analysis test of independence was conducted to determine if the number of grade levels a teacher taught was related to their students' reading level; however these analyses were non-significant for above-grade, χ^2 (4, N = 314) = 2.69, p > .05; on-grade χ^2 (4, N = 311) = 4.05, p > .05, and below-grade readers, χ^2 (4, N = 311) = 7.71, p > .05. Therefore, there was no relationship between the grade levels a teacher taught and the reading level of their students. Thus, although grade level was a significant contributor to teacher self-efficacy on the STERI, this difference could be due to the grouping of teachers into single or multiple grade levels taught.

Percent of below-grade readers. A one-way ANOVA was conducted to determine if the percent of below-grade readers teachers reported in their classes was a significant contributor to teachers' reported scores on the STERI. Levene's Homogeneity of Variance was violated (p < .05), and results of the ANOVA were significant, F(4, 275) = 3.73, p < .01, η^2 = 5.43%. Because Levene's Homogeneity of Variance was violated (p < .05), a Games-Howell post-hoc was conducted to determine where the significant difference were between groups. Teachers with 75% or more below-grade readers had a statistically significantly higher mean on the STERI than those who taught 0% (M_{diff} = 1.48, p < .01, d = 1.56), 1-25% (M_{diff} = 1.29, p < .01, d = 1.14), or 26-50% (M_{diff} = 1.10, p < .05, d = 1.42) of below-grade readers. Therefore, teachers who had a higher percentage of below-grade readers felt more confident in their abilities in literacy instruction than those teachers who had lower percentages of below-grade readers.

Title I status. When examining if teaching in a Title I school affected teachers' scores on the STERI an independent t-test was conducted. Levene's Homogeneity of Variance was violated (p < .05), and significant difference were reported, t(140.27) = -

4.43, $p < .001$, $d = .58$, with teachers teaching at Title I scoring higher ($M_{\text{diff}} = .59$) on the STERI than those who did not. This finding contradicts Tschannen-Moran and Hoy's (2007) finding that school setting does not have a relationship to teachers' self-efficacy. To examine this difference further, a chi-square test of independence was conducted on Title I status and each of the three types of readers (i.e., above-grade, on-grade, and below-grade). Chi-square analysis revealed a significant relationship when examining Title I status in conjunction with percent of on-grade readers, χ^2 (4, N = 314) = 13.61, $p < .01$ and below-grade readers, χ^2 (4, N = 311) = 12.30, $p < .05$. Therefore, there was a relationship between teaching in a Title I school and teaching both on and below-grade readers. This finding may be explained by the increased number of below-grade readers that are enrolled in Title I schools. Further, it is not surprising that Title I teachers reported a higher mean score on the STERI when coupled with the relationship between Title I schools and percent of on and below-grade readers. Hence, as Title I teachers have higher percentages of on and below-grade readers, they have higher self-efficacy in reading instruction at the secondary level.

Content area. To assess if content area made a difference in teachers' reported self-efficacy on the STERI, a one-way ANOVA was conducted. The overall model was significant, $F(3, 282) = 88.03$, $p < .001$, $\eta^2 =$ 93.65%, thus there was statistically significant difference based on teachers' content area. As Levene's Homogeneity of Variance was violated ($p < .01$), to examine differences a Games-Howell post-hoc was conducted. The post-hoc analysis revealed that there was a statistically significant difference between each of the four content areas. More specifically, English teachers reported a higher level of self-efficacy on the overall STERI than teachers in social studies (M_{diff} = .61, $p < .001$, d = .87), science (M_{diff} = 1.46, $p < .001$, d = 1.89), and mathematics (M_{diff} = 1.86, $p < .001$, d = 2.28). Further, social studies teachers reported a higher mean than science (M_{diff} = .85, $p < .001$, d = 1.05) and mathematics (M_{diff} = 1.24, $p < .001$, d = 1.47) teachers. See Table 31 for full breakdown of differences by content area. These differences partially align with the hypotheses made in Chapter 3, that teachers of English would have the highest self-efficacy, followed by social studies, science, and then mathematics based on the volume of text and readings conducted in each content area.

Table 31

Group Mean Differences Based on Teachers' Content Area

	Overall STERI	Using Reading Strategies	Motivating Student to Read	Teaching Struggling Readers
Content Area	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
English	4.62 (0.67)[a]	4.67 (0.68)[a]	4.80 (0.61)[a]	4.49 (0.78)[a]
Social Studies	4.01 (0.75)[b]	3.55 (1.06)[b]	4.48 (0.84)[a]	3.99 (0.93)[b]
Science	3.16 (0.87)[c]	2.88 (1.07)[c]	3.60 (0.86)[b]	3.06 (1.01)[c]
Mathematics	2.77 (0.94)[c]	2.66 (1.22)[d]	2.98 (.97)[b]	2.60 (0.89)[c]

Note. Means sharing a letter in their superscript are not significantly different at the .05 level according to a Games-Howell post-hoc test.

Summary of contextual factors. One-way ANOVA, independent *t*- tests, and chi-square analyses were conducted to examine if particular teacher and environmental factors were related to teachers' reported self-efficacy on the STERI. Teachers factors that were examined were: gender, years of experience, highest level of education, preparation (i.e., number of literacy course taken for credit and number of professional development courses), and ethnicity/race. Aligning with previous self-efficacy research (e.g., Raudenbush et al., 1992), teachers' years of experience and ethnicity/race did not relate to their scores. Additionally, both gender and preparation were found to significantly contribute to teachers' scores on the STE6RI; affirming previous findings in the literature (e.g., Klassen & Chui, 2010; Tschannen-Moran & Johnson, 2011). Further, level of education was found to significantly relate to teachers' scores. Therefore, further analyses must be conducted with a new sample to determine if level of education contributes to teachers' scores on the STERI.

In addition, environmental factors were also analyzed to determine if they were related to teachers' scores on the STERI. The environmental factors examined were: number of courses taught, grade level taught, Title I school status, percent of students that were ELLs, percent of students with IEPs, students' reading level, and teachers' primary content area of instruction. Of these variables, number of courses taught, percent of students that were ELLs, percent of students with IEPs, and percent of on-grade and above-grade readers did not significantly contribute to teachers' STERI scores. Moreover, two significant factors that contributed to teachers' scores, were percent of below-grade readers and Title I school status. Interestingly, chi-square analyses revealed a statistically significant relationship between the two variables. Further, teachers in Title I schools scored higher on the STERI than non-Title I teachers. One assumption to support this relationship is that Title I schools have more below-grade readers and therefore teachers at these schools have higher self-efficacy for reading instruction because they are engaged in these practices at the secondary level.

Upon initial examination of the data, certain teacher and environmental factors contributed to teachers' scores on the STERI. However, these variables showed significance for this particular sample. Therefore, future research must examine these factors to determine if their relationships are consistent, or just a reflection of sampling. Although these factors may only be significant for this sample, the notion of examining teacher self-efficacy in relation to students' reading level is new to the field. As no other study has examined teacher self-efficacy in relation to students reading level, these findings contribute to the field in a new and meaningful way.

Reliability

Reliability examines the consistency of survey responses (Crocker & Algina, 1986). In this study, internal consistency reliability was reported using Cronbach's alpha (α; Cronbach, 1951). Reliability is generally reported for each factor within an instrument with .70 - .90 considered good, and .90 or higher as excellent (Cortina, 1993; DeVellis, 2003). The data for each of the three subscales on the STERI were found to have excellent internal consistency reliability and high inter-item correlations. Data on Factor 1, self-efficacy for Using Content-Area Reading Strategy Instruction had the highest reliability ($\alpha = .96$) with inter-item correlations ranging from .72 - .86. On Factor

2, self-efficacy for Motivating Students to Read, data had excellent internal consistency (α = .94) with inter-item correlations ranging from .64 - .83. Finally, on Factor 3, self-efficacy for Teaching Struggling Readers, data had excellent reliability (α = .95) and inter-item correlations with a range of .59 - .86. In addition, when examining each subscale it was found that the deletion of any item would decrease Cronbach's alpha; providing further support for the inclusion of the 22 items on the STERI, with an overall reliability (α = .98). Based on Cronbach's alpha measure of internal consistency, the data have excellent reliability across all three subscales of the STERI.

Once reliability was established mean composites were created for each of the three STERI subscales: self-efficacy for Using Content-Area Reading Strategy Instruction, self-efficacy for Motivating Students to Read, and self-efficacy for Teaching Struggling Readers. Mean composites were created using the same process as the convergent and discriminant scales (i.e., item scores for each subscale were summed and then the total score was divided by the number of items). For example, for the self-efficacy for Using Content-Area Reading Strategy Instruction subscale, the composite was calculated by adding the raw score of the seven items and dividing that total by seven for each individual. These mean composites were used in correlational analyses to determine the validity of the STERI.

Validity

Convergent validity establishes evidence of similarity of measurement between theoretically related constructs (DeVellis, 2003; Kerlinger, 1973) and discriminant validity tests if an instrument was measuring something it did not intend (DeVellis, 2003). In this study, to establish convergent and discriminant validity, correlational analysis using Pearson's *r* was conducted for each of the subscales of the STERI and the subscales of the TSES, GOT, and the Invitations of Self/Others Scale (see Table 32). For the purposes of this study, strengths of relationships were addressed as: a) very weak correlation (r = .20 or less); b) weak correlation (r = .20 - .40); c) moderate correlation (r = .50 - .70); and d) strong correlation (r = .80 and higher).

Table 32

Correlations Among STERI Subscales, TSES, GOT, and Invitations to Self/Others Subscales

	1	2	3	4	5	6	7	8
1. STERI Teaching Reading Strategies								
2. STERI Motivating Students to Read	.73**							
3. STERI Struggling Readers	.81**	.82**						
4. TSES Engagement	.25**	.43**	.35**					
5. TSES Strategy Instruction	.50**	.34**	.40**	.51**				
6. GOT Mastery Goals	.25**	.25**	.21**	.17**	.25**			
7. GOT Relational Goals	.29**	.30**	.25**	.27**	.27**	.33**		
8. Invitations Self	.11	.33**	.27**	.36**	.14*	.22**	.13*	
9. Invitations Others	.43**	.28**	.27**	.30**	.49**	.26**	.36**	.18**

Note. * = *p* < .05. ** = *p* < .001. STERI = Secondary Teachers' Self-Efficacy for Reading Instruction. TSES = Teachers' Sense of Efficacy Scale. GOT = Goal Orientations for Teaching.

Correlations among STERI subscales. It was hypothesized that all three subscales would be weakly to moderately positively correlated as they all dealt with secondary reading instruction (see Chapter 3 for detailed hypotheses). However, the relationships should have only been weak to moderate as each of the subscales represented a different aspect of reading instruction at the secondary level. Additionally, the strongest positive correlation was hypothesized to be between Factor 1 (Using Content-Area Reading Strategy Instruction) and Factor 3 (Teaching Struggling Readers), as these two tasks often occur in tandem. Factor 1 (Using Content-Area Reading Instruction) and Factor 2 (Motivating Students to Read) were hypothesized to have the weakest positive relationship as teachers may not have interpreted these concepts as overlapping. Finally, Factor 2 (Motivating Students to Read) and Factor 3 (Teaching Struggling Readers factor) were hypothesized to be weakly positively correlated as teachers may view struggling readers as unable to be motivated.

The first hypothesis that all three subscales would be positively weakly to moderately correlated was partially upheld: Factor 1 with Factor 2 ($r = .73$, $p < .01$), Factor 1 with Factor 3 ($r = .81$, $p < .01$), and Factor 2 with Factor 3 ($r = .82$, $p < .01$). These correlations revealed positively moderate to strong relationship between each of the three factors of the STERI; however, these relationships were stronger than originally anticipated. The hypothesis of the strongest relationship between Factors 1 and 3 was nearly confirmed as these two factors had the second highest correlation ($r = .81$, $p < .01$). Yet, this relationship was as expected in that teachers who had a higher sense of efficacy in teaching reading strategies would also have a higher sense of efficacy in teaching struggling readers; as the use of reading strategies can be compounded by the number of struggling readers a teacher teaches.

Additionally, the second hypothesis that the Using Content-Area Reading Strategy Instruction (Factor 1) and Motivating Students to Read (Factor 2) factors would have the weakest relationship was upheld as these two subscales had the lowest correlation (r = .73, p < .01). This relationship was as expected because teachers may not respond that reading strategies and motivation to read are overlapping variables. The final hypothesis that the Motivating Students to Read factor (Factor 2) would be weakly positively correlated with the Teaching Struggling Readers factor (Factor 3) was refuted as these two subscales had the highest positive correlation (r = .82, p < .01). This relationship contradicted the idea that teachers would not respond that teaching struggling readers and motivating students to read were overlapping constructs. Interestingly, teachers in this sample did draw the connection between these two concepts creating a highly positive correlation.

Convergent validity. To assess convergent validity, correlations were conducted between the subscales of the STERI (i.e., self-efficacy for Using Content-Area Reading Strategy Instruction, Motivating Students to Read, Teaching Struggling Readers), the subscales of the TSES (i.e., self-efficacy for Instructional Strategies, Engagement; Tschannen Moran & Woolfolk Hoy, 2001), and the subscales of the GOT (i.e., Mastery Goals, Relational Goals; Butler 2012, Butler & Shibaz, 2014).

TSES. In Chapter 3 hypotheses were made that the Using Content-Area Reading Strategy Instruction factor on the STERI would be moderately positively related to the both the TSES Instructional Strategies factor and the TSES Engagement factor; with the stronger relationship between the reading strategies and instructional strategies subscales. In regards to the STERI factor of Motivating Students to Read, hypotheses were made that the subscale would moderately positively correlate with both of the TSES subscales; with the stronger of the two relationship occurring between the STERI factor and the TSES Engagement subscale. Finally, it was hypothesized that the STERI Teaching Struggling Readers factor would be weakly positively correlated with both the TSES Instructional Strategies and Engagement factors.

Using Content-Area Reading Strategy Instruction factor. First, the STERI Using Content-Area Reading Strategy Instruction factor was weakly positively correlated to the TSES Engagement factor. Although this relationship was weak (r = .25, p < .01), it was a statistically significant correlation and demonstrated that these two factors were related, but still two different constructs. This supports the notion that the STERI factor of Using Content-Area Reading Strategy Instruction was theoretically different than the TSES Engagement factor. The second hypothesis was confirmed as the correlation between the STERI Using Content-Area Reading Strategy Instruction and the TSES Instructional Strategies was a significant moderate positive correlation (r = .50, p < .01). Thus the two factors are measuring similar constructs, however the moderate correlation demonstrates that they are not too similar (i.e., they are not measuring the same construct in the exact same way).

Motivating Students to Read factor. Correlational analyses demonstrated a weak positive relationship between STERI factor of Motivating Students to Read and the TSES

Instructional Strategies subscale (r = .34, p < .01). This partially supported the initial hypothesis as the relationship was positive, however it was weaker than originally hypothesized. This correlation shows that the subscales are significantly related, meaning they measure a similar construct, however they are distinct measures, as the correlations were not too strong. The correlation between the STERI factor Motivating Students to Read and the TSES Engagement factor was significantly moderately positive (r = .43, p < .01). These analyses partially support the initial hypothesis as the relationship between the STERI Motivating Students to Read and TSES Engagement factor would be stronger than the relationship between the STERI Motivating Students to Read and the TSES Instructional Strategies factor. However, both of these relationships were weakly to moderately correlated, and not as strong as hypothesized. The correlations between the subscales demonstrate that the STERI factor of Motivating Students to Read was related to the two TSES factors, but not so much as to imply the constructs were completely overlapping.

Teaching Struggling Readers factor. The STERI Teaching Struggling Readers factor was significantly weakly positively related to the TSES Instructional Strategies factor (r = .40, p < .01) and the TSES Engagement factor (r = .35, p < .01); supporting the original hypotheses. Thus, the STERI Teaching Struggling Readers factor was related to both the TSES factors, however neither of these relationships was too highly correlated to suggest issues of overlap in measurement. Meaning the Teaching Struggling Readers factor on the STERI was measuring a similar construct as the TSES subscales, but in a distinct way.

Summary. When examining the STERI subscales of: Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers in relation to the TSES Instructional Strategies and Engagement factors, correlational analyses showed all relationships were significant, positive, and weak to moderate (r = .25 - .50, ps < .01). These analyses demonstrated that the STERI was measuring a similar construct as the TSES subscales, however in a theoretically distinct manner. As none of the correlations were higher than r = .50, the analysis showed that there was slight overlap in measurement between the STERI and the TSES subscales, but nothing significant enough to warrant concern. Thus, convergent validity was established as the relationships between all subscales were significant, but not too highly correlated to suggest total overlap in measurement. Therefore, the STERI was measuring the intended construct (teacher self-efficacy) in a different way than the TSES.

GOT. Hypotheses in Chapter 3 predicted that the STERI factor of Using Content-Area Reading Strategy Instruction would be weakly positively related to both the GOT Mastery and Relational factors. In addition, it was hypothesized the STERI factor of Motivating Students to Read would be moderately positively related to the GOT Mastery factor and weakly positively related to the GOT Relational factor. Finally, it was predicted that the STERI Teaching Struggling Readers factor would be weakly positively related to the GOT Mastery factor and moderately positively related to the GOT Relational factor.

Using Content-Area Reading Strategy Instruction factor. Correlational analyses revealed that the STERI Using Content-Area Reading Strategy Instruction factor was significantly weakly positively correlated with the GOT Mastery ($r = .25$, $p < .01$) and the GOT Relational ($r = .29$, $p < .01$) factors. This supported the hypotheses made in Chapter 3 and demonstrated that the STERI Using Content-Area Reading Strategy Instruction factor was related to the GOT subscales, but in a theoretically different way as both correlations were weak. Additionally, the relationship between the STERI Using Content-Area Reading Strategy Instruction factor and the GOT factors mirrored the correlations between the TSES Instructional factor and the GOT Mastery ($r = .25$, $p < .01$) and Relational ($r = .27$, $p < .01$) factors. These correlations further established convergent validity as the STERI reading strategy subscale was related to the GOT factors in a similar manner to the TSES subscales.

Motivating Students to Read factor. The STERI Motivating Students to Read factor was significantly weakly positively correlated to the GOT Mastery ($r = .25$, $p < .01$) and Relational factors ($r = .30$, $p < .01$). This partially supported the hypotheses made in Chapter 3, however the relationship between the STERI Motivating Students to Read factor and the GOT Mastery factor was weaker than predicted. Although the hypotheses were not fully supported, the relationship between these factors show that the STERI subscale and the GOT subscales are statistically significantly related, but not more than weakly. Convergent validity was further established when examining the correlations between the TSES Engagement factor and the GOT Mastery ($r = .17$, $p < .01$) and Relational ($r = .27$, $p < .01$) factors. These corrections followed a similar pattern as the correlations between the STERI subscale and the GOT subscales. More specifically, the STERI Motivating Students to Read factor was related to the GOT factors in the same way as the TSES Engagement factor.

Teaching Struggling Readers factor. Correlational analyses revealed that the STERI Teaching Struggling Readers factor was weakly positively related to both the GOT Mastery (r = .21, p < .01) and Relational (r = .25, p < .01) factors. These results partially supported the hypotheses made in Chapter 3, however the relationship between the STERI Teaching Struggling Readers factor and the GOT Relational factor was not as strong as predicted. Although the hypotheses were not fully confirmed, the statistically significant weak relationships demonstrated that the STERI Teacher Struggling Readers factor was somewhat related to the GOT Mastery and Relational subscales.

Summary. Convergent validity was established by examining the correlations between the STERI factors of Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers with the GOT Mastery and Relational factors. Each of these correlations revealed that the STERI subscales were statistically significantly weakly related to the GOT subscales (r = .17 – r = .30, ps < .01). Therefore, the STERI subscales were related to the GOT subscales, but in different ways; as shown by the weak correlations. Furthermore, the weak relationships between the STERI subscales and the GOT subscales provided support that the STERI is not measuring the same construct as the GOT. Thus, the GOT provided further convergent validity that the STERI is measuring teacher's self-efficacy and not goal orientations.

Summary of convergent validity. Convergent validity was initially established through correlational analyses between the STERI subscales: Using Content-Area

Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers and the TSES subscales: Instructional Strategies and Engagement. These correlations showed significant weak to moderate positive ($r = .25$ - $r = .50$, ps $< .01$) relationships between the subscales; demonstrating that the subscales were measuring similar constructs, but in distinct ways. Further, convergent validity was shown by examining the aforementioned STERI subscales and the GOT factors of Mastery and Relational Goal Orientations. The significant weak correlations ($r = .21$ - $r = .30$, ps $<$ .01) between these subscales demonstrated that the constructs were somewhat related, but with no significant overlap in measurement.

Additionally, convergent validity was established when examining the patterns in relationships between the GOT Mastery and Relational factors with both the TSES Instructional Strategies and Engagement factors and the STERI Using Content-Area Reading Strategy Instruction and Motivating Students to Read factors. As the TSES and STERI subscales were related (see above), the matching patterns in their correlations with the GOT subscales further established convergent validity. Thus, the correlational analyses established convergent validity across the STERI subscales and demonstrated that the STERI was measuring teachers' self-efficacy, but without excessive overlap in measurement. In particular, by not having multicollinearity with these instruments (i.e., no correlations were above .50), the STERI assessed self-efficacy in a theoretically similar manner as preexisting instruments (i.e., TSES), but with distinct measurement.

Discrimination from TSES Scales. Although moderate correlations provided convergent validity, the STERI functioned differently than the TSES when means, ranges, ANOVAs, and correlations were conducted on the data separated by content area (i.e., science, social studies, mathematics, English). First, mean composites revealed that a pattern was present in teachers' sense of efficacy on each of the STERI subscales (see Table 31). The pattern of highest to lowest mean on all three STERI subscales was: English, social studies, science, and mathematics; confirming hypotheses from Chapter 3. On each of the three STERI subscales the same pattern arose with English teachers having the highest mean, followed by social studies, then science, and finally mathematics. Therefore, the STERI was able to discriminate based on content area, therefore content area made a different when examining teachers' mean scores on the STERI. This type of pattern did not occur with the TSES subscales of Engagement and Strategy Instruction. For the TSES Engagement factor, social studies teachers had the highest mean, followed by English, then mathematics, and science. For the TSES Strategy Instruction factor, English teachers had the highest mean, followed by social studies, then mathematics, and finally science.

Further, the STERI items were measured on a six-point scale, however the only teachers to show the full range (i.e., 1-6) on any of the STERI subscales were English teachers; all other content area teachers did not have a maximum range of six (see Table 33). Therefore, teachers in social studies, science, and mathematics were not showing any type of ceiling effect and teachers' responses were varied. This pattern does not exist for the TSES subscales where all teachers had ranges from 1-9 (i.e., the complete scale), thus teachers were responding to the highest anchor on the scale. This suggests that the

STERI discriminated across content areas differently than the TSES; a finding that was further explored through the use of one-way ANOVAs.

Table 33

Descriptive Statistics for STERI and TSES Subscales by Teachers' Content Area

	Min.	Max.	Range	Mean	St. Dev.
English					
STERI Teaching Reading Strategies	1.57	6.00	4.43	4.67	0.78
STERI Motivating Students to Read	2.00	6.00	4.00	4.80	0.61
STERI Struggling Readers	1.38	6.00	4.63	4.49	0.78
TSES Engagement	1.00	9.00	8.00	6.43	1.31
TSES Strategy Instruction	1.25	9.00	7.75	7.33	1.14
Social Studies					
STERI Teaching Reading Strategies	2.00	5.71	3.71	3.55	1.06
STERI Motivating Students to Read	2.25	5.88	3.63	4.48	0.84
STERI Struggling Readers	1.25	5.38	4.13	3.99	0.93
TSES Engagement	3.00	9.00	6.00	6.73	1.31
TSES Strategy Instruction	3.75	9.00	5.25	6.80	1.38
Science					
STERI Teaching Reading Strategies	1.00	5.43	4.43	2.88	1.07
STERI Motivating Students to Read	2.00	5.50	3.50	3.60	0.86
STERI Struggling Readers	1.00	5.38	4.38	3.06	1.01
TSES Engagement	2.50	9.00	6.50	5.92	1.27
TSES Strategy Instruction	2.75	9.00	6.25	6.38	1.52
Mathematics					
STERI Teaching Reading Strategies	1.00	5.14	4.14	2.66	1.22
STERI Motivating Students to Read	1.00	5.00	4.00	2.98	0.97
STERI Struggling Readers	1.00	5.25	4.25	2.60	0.89
TSES Engagement	1.75	8.50	6.75	6.33	1.41
TSES Strategy Instruction	2.00	9.00	7.00	6.75	1.46

Note. Min. = Minimum. Max. = Maximum. St. Dev. = Standard Deviation. STERI = Secondary Teachers' Self-Efficacy for Reading Instruction. TSES = Teachers Sense of Efficacy Scale.

Mean differences. Five one-way ANOVAs were conducted (one for each of the three STERI subscales and the two TSES subscales) to determine if the mean differences in scores were statistically different based on content area (see Table 31). For the three STERI subscales (i.e., Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers) all three ANOVAs showed statistically significant differences across content areas in a similar pattern. The Using Content-Area Reading Strategy Instruction factor showed significant difference across content areas, F (3, 298) = 75.47, $p < .001$, $\eta^2 = 43.17\%$. A Games-Howell post-hoc was examined as Levene's Homogeneity of Variance was violated ($p < .001$) to determine where the significant different were amongst content area teachers (see Table 31). English teachers had a statistically significant ($p < .001$) higher mean than social studies ($M_{\text{diff}} = 1.12$, $d = 1.20$), science ($M_{\text{diff}} = 1.78$, $d = 1.92$), and mathematics teachers ($M_{\text{diff}} = 2.00$, $d = 1.96$). Social Studies teachers had a statistically significantly ($p < .01$) higher mean than science ($M_{\text{diff}} = .66$, $d = .63$) and mathematics teachers ($M_{\text{diff}} = .88$, $d = .78$). Finally, there was no significant difference between science and mathematics teachers ($p > .05$). Therefore, the pattern for the Using Content-Area Reading Strategy Instruction subscale was English, social studies, science, and mathematics with highest to lowest mean difference. This supported the hypotheses made in Chapter 2 that English teachers would have the highest mean followed by social studies, then science, and mathematics.

In addition, the STERI Motivating Students to Read factor also had statistically significant differences across content areas, F(3, 301) = 81.44, $p < .001$, $\eta^2 = 44.80\%$. Post-hoc analyses were conducted using Games-Howell as Levene's Homogeneity of

Variance was violated ($p < .001$). These analyses showed that English teachers had a statistically significantly ($p < .001$) higher mean than both science ($M_{\text{diff}} = 1.20$, $d = 1.61$) and mathematics teachers ($M_{\text{diff}} = 1.81$, $d = 2.22$), but not social studies teachers ($p > .05$; see Table 31). Social studies teachers had a statistically significantly ($p < .001$) higher mean difference than science ($M_{\text{diff}} = .88$, $d = 1.04$) and mathematics teachers ($M_{\text{diff}} = 1.49$, $d = 1.65$). Finally, science teachers had a statistically significantly ($p < .01$) higher mean than mathematics teachers ($M_{\text{diff}} = .62$, $d = .67$). These data follow a similar to pattern to the means on the STERI Using Content-Area Reading Strategy Instruction factor with the pattern of highest to lowest means: English, social studies, science, and mathematics.

Further, the STERI Teaching Struggling Readers factor showed statistically significant differences across content areas, F (3, 300) = 74.53, $p < .001$, $\eta^2 = 42.70\%$. To examine these differences, a Games-Howell post-hoc analysis was conducted as Levene's Homogeneity of Variance was violated ($p < .05$). Following a similar pattern to the two other subscales of the STERI, English teachers had a statistically significantly ($p < .01$) higher mean than social studies ($M_{\text{diff}} = .50$, $d = .58$), science ($M_{\text{diff}} = 1.43$, $d = 1.58$), and mathematics teachers ($M_{\text{diff}} = 1.89$, $d = 2.26$). Social studies teachers had a statistically significantly ($p < .001$) higher mean than science ($M_{\text{diff}} = .93$, $d = .96$) and mathematics teachers ($M_{\text{diff}} = 1.39$, $d = 1.53$). And science teachers had a statistically significantly ($p < .05$) higher mean than mathematics teachers ($M_{\text{diff}} = .46$, $d = .48$). Just as the other two subscales of the STERI the highest to lowest means followed the same pattern: English, social studies, science, and mathematics.

When examining the TSES subscales, patterns were not seen in such an explainable manner. The TSES Instructional Strategies factor was found to be statistically significantly different across content areas, $F(3, 299) = 8.32$, $p < .001$, $\eta^2 = 7.7\%$. Because Levene's Homogeneity of Variance was violated ($p < .05$), a Games-Howell post-hoc was conducted to determine which content areas significantly differed. Post-Hoc analysis revealed that English teachers had a statistically ($p < .05$) higher mean score than both science ($M_{\text{diff}} = .95$, $d = .71$) and mathematics ($M_{\text{diff}} = .56$, $d = .44$) teachers. However, it is assumed that this pattern is only evident because of the presentation of the overall survey package. Participants took the survey in the following order: STERI, TSES, GOT, and Invitations scales. As teachers had already read reading strategy instruction items in the STERI, they may have answered the TSES instructional strategies items in the mindset of reading strategies. Therefore, the difference in scores may be attributed to the order of the survey package and further studies should address this by varying the order of the survey instruments.

For the TSES Engagement factor, the overall ANOVA was statistically significant $F(3, 303) = 4.59$, $p < .01$, $\eta^2 = 4.34\%$, thus differences occurred across content areas. To explore these differences, Tukey HSD post-hoc was conducted as Levene's Homogeneity of Variance was not violated $p > .05$. Results of the post-hoc revealed that English teachers had a statistically significantly ($p < .05$) higher mean than science teachers ($M_{\text{diff}} = .50$, $d = .40$). And social studies teachers had a higher mean than science teachers ($M_{\text{diff}} = .80$, $d = .63$). All other content areas showed no significant mean differences on the TSES Engagement factor.

Therefore, based upon the patterns in mean scores the STERI subscales all function in similar ways across content areas; whereas the TSES subscales did not follow a similar pattern across content areas. This provided further evidence that the STERI subscales are stable and discriminate across content areas in a similar and theoretically expected pattern; with means from high to low: English, social studies, science, and mathematics. The TSES did not follow this pattern and did not statistically significantly discriminate across all content areas. Therefore, the STERI measured self-efficacy in a different way than the TSES; providing further evidence that although both scales measure teacher self-efficacy, they do so in theoretically different ways.

Correlational differences. Additionally, correlations among the subscales were examined by content area to determine if teachers of different content areas viewed the relationships of the subscales of the STERI and the TSES differently (see Table 34). When examining the difference between the correlations of the STERI subscales and TSES subscales one important finding was that teachers across all content areas reported the three STERI subscales to be correlated. English, social studies, science, and mathematics teachers all responded that the STERI subscales were weakly to highly positively correlated ($r = .35$ – $r = .80$; ps < .01). Suggesting that teachers of all content areas responded to the items of the STERI in a similar way; refuting a hypothesis made in Chapter 3 that the correlations among subscales would be different based on content area.

However, the consistent relational pattern reported by teachers demonstrated that the subscales of the STERI functioned the same across content areas (i.e., all are statistically significantly related), and discriminated mean scores reported by each content area (i.e., mean differences between English, social studies, science, and mathematics). More specifically, the STERI subscales were responded to consistently across content areas, showing stability among each subscale. Further, the ability of the subscales to show statistically significant mean differences amongst content areas revealed that even though teachers across content areas responded to the subscales similarly, they answered items differently. Interestingly, the teachers who responded that the subscales the most highly correlated were mathematics teachers (r = .71 - r = .79, ps < .01); yet these were also the teachers with the lowest means on the STERI subscales. This suggested mathematics teachers responded that reading strategy instruction, motivating students to reading, and teaching struggling readers were very similar tasks and that is why they reported lower means on all three subscales.

The correlations for the TSES subscales were not consistent across content areas, and findings were not as easily interpreted as the STERI. English, science, and mathematics teachers responded that the subscales (i.e., Instructional Strategies and Engagement) moderately positively correlated (r = .53 - r = .72; ps < .001); however social studies teachers responded the relationship weak and non-significant (r = .21, p > .05). There is no logical interpretation of this finding, other than that the TSES subscales did not demonstrate patterns in mean scores across content areas, so it is understandable that teachers viewed their correlations differently.

Table 34

Correlation of Subscales by Teachers' Content Area

	1	2	3	4
English				
1. Teaching Reading Strategies				
2. Motivating Students to Read	.64**			
3. Helping Struggling Readers	.76**	.70**		
4. TSES Engagement	.42**	.58**	.48**	
5. TSES Instructional Strategies	.46**	.62**	.61**	.60**
Social Studies				
1. Teaching Reading Strategies				
2. Motivating Students to Read	.35**			
3. Helping Struggling Readers	.54**	.61**		
4. TSES Engagement	.00	.66**	.32*	
5. TSES Instructional Strategies	.57**	.21	.28*	.21
Science				
1. Teaching Reading Strategies				
2. Motivating Students to Read	.62**			
3. Helping Struggling Readers	.70**	.67**		
4. TSES Engagement	.31**	.49**	.36**	
5. TSES Instructional Strategies	.53**	.17	.19	.53**
Mathematics				
1. Teaching Reading Strategies				
2. Motivating Students to Read	.71**			
3. Helping Struggling Readers	.75**	.79**		
4. TSES Engagement	.19	.22	.27*	
5. TSES Instructional Strategies	.27*	.14	.25	.72**

Note. * = $p < .05$. ** = $p < .001$.

Therefore, the examination of means, ranges, one-way ANOVAs, and correlations demonstrated that although the STERI functioned similarly to the TSES when examining convergent validity, it discriminated across content areas in more predictable and logical ways than the TSES. More specifically, the STERIs ability to discriminate across content areas in a logical manner demonstrates that the STERI does measure teachers' self-efficacy in a distinct way from the TSES. Further, item means and ranges showed that the STERI had items that did not lead to ceiling effects (i.e., only English teachers reported scores on the maximum range), whereas the TSES had ceiling effects (i.e., reporting of the highest maximum score) across all four content areas. Thus, the STERI and the TSES both measured teachers' self-efficacy, however they do so in theoretically distinct ways, providing rationale for the inclusion of the STERI as an instrument to measure teachers' sense of efficacy in reading instruction across the content areas.

Discriminant validity. Discriminant validity assesses if an instrument is measuring something it was not intended to measure. Therefore, in order for discriminant validity to be established, the new measure should be uncorrelated with preexisting validated instruments that measure unrelated constructs. Discriminant validity was partially established by conducting correlational analyses between the subscales of the STERI and the subscales of Invitations to Self/Others (Morris, 2010).

Invitations to Self/Others. The Invitations to Self/Others is a scale that contained two subscales: invitational messages a teacher sends to himself (Self), and the messages one sends to other teachers (Others). In Chapter 3 hypotheses predicted that the Invitations to Self subscale would not be significantly correlated to any of the STERI subscales. In addition, it was hypothesized that the Invitations to Others subscale would be weakly positively correlated with the STERI subscales.

Invitations to Self factor. Correlational analyses revealed that the Invitations to Others subscale was significantly weakly correlated with the STERI Motivating Students to Read (r = .33, p < .01) and Teaching Struggling Readers (r = .27, p < .01) factors; however it was not significantly related to the STERI Using Content-Area Reading Strategy Instruction factor (r = .11, p > .05). These correlational analyses revealed that the STERI Motivating Students to Read and Teaching Struggling Readers factors were related to the Invitations to Others subscale, but not related to the STERI Using Content-Area Reading Strategy Instruction factor. Therefore, only one of the hypotheses was accepted.

Interestingly, a similar pattern arose when examining the relationship between the Invitations to Self factor and the TSES Engagement (r = .36, p < .01) and Instructional Strategies (r = .14, p < .05) factors. These analyses reveal that the STERI factors are related to the Invitations to Self subscale in the same manner as the TSES factors. Thus, although the analyses did not provide discriminant validity for the STERI Motivating Students to Read and Teaching Struggling Readers factors; they did provide discriminant validity for the STERI reading strategies factor and further convergent validity for the STERI Motivating Students to Read factor when examining the patterns with the TSES subscales.

Invitations to Others factor. Similarly to the Invitations to Self subscale, correlational analyses revealed the Invitations to Others subscale was significantly weakly to moderately positively correlated to the STERI Unitizing Content-Area Reading Strategy Instruction ($r = .43$, $p < .01$), Motivating Students to Read ($r = .28$, $p < .01$), and Teaching Struggling Readers ($r = .27$, $p < .01$) factors. These results partially supported the hypotheses that the Invitations to Others subscale would be weakly correlated with the STERI subscales. Although this hypothesis was only partially supported, just as with the Invitations to Self subscale, the Invitations to Others subscale correlated with the TSES factors in a similar pattern. The Invitations to Others subscale significantly positively weakly to moderately correlated with the TSES Instructional Strategies ($r = .50$, $p < .01$) and Engagement ($r = .30$, $p < .01$) factors. Therefore, the STERI factors correlated with the Invitations to Others subscale in the same pattern as the TSES subscales; providing further convergent validity for the STERI. Thus, even though the hypotheses were not fully supported, the pattern of the STERI Using Content-Area Reading Strategy Instruction factor and TSES Instructional Strategies factor following a similar pattern in their relationship with the Invitations to Other subscales provides further convergent validity for the STERI subscale.

Summary of discriminant validity. The Invitations to Self subscale provided partial discriminant validity for the STERI reading strategies factor. Whereas the Invitations to Self subscale provided better discrimination with the STERI Using Content-Area Reading Strategy Instruction factor, it was less discriminating with the STERI Teaching Struggling Readers factor. Although this relationship was weak ($r = .27$, $p < .01$), the significance of this relationship did not provide a clear measure of discrimination. However, the Invitations to Self subscale provided further convergent validity for the STERI Motivating Students to Read subscale as the correlational pattern for this factor was similar to the TSES Engagement factor.

Additionally, the Invitations to Others subscale functioned much like the Invitations to Self subscale in that it did not provide clear discrimination among the STERI factors. However, it provided further convergent validity for the STERI Using Content-Area Reading Strategy Instruction factor as the relationship between this factor and the Invitations to Others subscale mirrored that of the TSES Instructional Strategies factor. Thus, partial discriminant validity was only established for the STERI Using Content-Area Reading Strategy Instruction factor and the Invitations to Self subscale; however, further convergent validity was established for the STERI Using Content-Area Reading Strategy Instruction and Motivating Students to Read factors.

Summary

The purpose of this study was to develop and validate an instrument to measure secondary teachers' sense of efficacy in reading instruction (STERI) across the content areas. This chapter described how experts in the fields of adolescent literacy and self-efficacy examined the STERI items for clarity and relevance. After revisions were made to the STERI items, cognitive interviewing was utilized to ensure that the STERI was a qualitatively sound instrument before engaging in exploratory analyses. Upon engaging in two rounds of cognitive interviews, the initial STERI items were revised and the final 31 item instrument was ready for exploratory analyses.

Following the seven step procedure outlined in Chapter 3, data were collected from 319 secondary content area teachers and used to conduct an EFA. After descriptive statistics and EFA were conducted, 22 of the original 31 items remained. The 22 items loaded onto the original three hypothesized factors (i.e., Using Content-Area Reading Strategy Instruction, Motivating Students to Read, Teaching Struggling Readers) and explained 74.98% of the variance equally across the three factors (Factor 1: 28.26%, Factor 2: 25.69%, Factor 3: 24.32%). Reliability was established using Cronbach's alpha, with all values greater than .90.

Analysis of teacher and environmental contextual factors were examined to determine their relationship to teachers' scores on the STERI. One-way ANOVAs, independent *t*- tests, and chi-square tests of independence analysis were conducted to examine contextual effects on teachers' STERI scores. Although certain teacher (i.e., gender, level of education, and preparation) and contextual (i.e., grade level taught, Title I school status, percent of below-grade readers, and content area) factors were found to statistically significantly contribute to teachers' scores; these analyses must be conducted on a new sample of participants to confirm or refute these initial findings.

Convergent validity was established for the STERI subscales through significant weak to moderate positive correlations with the TSES and GOT subscales. Significant differences were revealed when examining teachers' primary content area and their overall score on the STERI, as well as their scores on each of the three STERI subscales. These differences confirmed hypotheses made in Chapter 3 that teachers' sense of self-efficacy on the STERI would follow a pattern of highest to lowest based on content area: English, social studies, science, and mathematics. Additionally, these findings provide discriminant validity from the TSES, as these patterns were not replicated when examining content area differences on the TSES. Further, discriminant validity was partially established for the STERI reading strategies subscale and the Invitations for Self subscale. In the following chapter I discuss these findings in greater detail, explain limitations to the current study, and provide directions for future research.

Chapter Five: Discussion

In this study I sought to develop and validate a teacher self-efficacy measure to capture Secondary Teachers' Sense of Efficacy for Reading Instruction (STERI) across the content areas. This chapter will describe the problem the STERI sought to address, provide a discussion of the conclusions, define limitations, offer implications for research and practice, and suggest recommendations for future research.

Statement of the Problem

With the rise of technology and digital texts, adolescent readers are faced with the challenges of engaging and interpreting print and nonprint texts to make meanings of their educational and social worlds (Biancarosa & Snow, 2007). Further, 62% of 12th grade students are reading at basic and below basic levels (Nations Report Card, 2013). As adolescents continue to struggle with reading, the necessity for educators to understand the difficulties and supports needed for student literacy development and success becomes paramount. Even though research has shown the importance for integrating literacy instruction at the secondary level, many content area teachers have negative beliefs about infusing literacy instruction in content area classrooms (Alger, 2007; Draper, 2002; Fisher & Ivey, 2005; Lesley, 2005). Given that teacher beliefs' affect instructional practices (Guo et al., 2010; Tschannen-Moran et al., 1997), it is important to measure these beliefs in relation to literacy instruction.

One specific belief that research has shown to be correlated with teacher instruction is teachers' beliefs about their effectiveness (self-efficacy) with a given task or certain context (Guo et al., 2014; Holzberger et al., 2013). If literacy instruction is a necessary component of secondary education, teachers must feel they can be effective; however, little is known about secondary teachers' self-efficacy for literacy instruction (Cantrell & Callaway, 2008; Cantrell & Hughes, 2008) as none of the current self-efficacy instruments focus on literacy instruction at the secondary level (Klassen et al., 2011). When researching secondary teachers' beliefs about literacy instruction in the secondary classroom, it is important to measure their self-efficacy in using these literary practices. Thus, the STERI was created to assess secondary teachers' sense of efficacy for reading instruction in their content classrooms.

Teacher beliefs and practice. Many secondary teachers believe literacy instruction is not their responsibility and think of themselves solely as content teachers, not teachers of reading (Fisher & Ivey, 2005; Jacobs, 2008; Kamil, 2003). Furthermore, research has shown that even when secondary teachers want to use content area literacy instruction in their classroom, they lack the training and resources necessary to effectively deliver literacy instruction (Gee & Forrester, 1988; Hall, 2005; Heller & Greenleaf, 2007; Sturtevant, 1996). Therefore, researchers must work to use teacher education and sustained professional development to combat the negative teacher beliefs surrounding content area literacy instruction and develop teachers' knowledge and skills related to literacy instruction in the secondary classrooms. Fisher and colleagues (2002) found that with sustained professional development structured around content area literacy instruction, teachers were able to incorporate these strategies into their classrooms, and students' reading achievement increased by two grade-level equivalents on a standardized reading assessment (i.e. Gates-MacGinitie).

As a way to foster secondary teachers beliefs and skills about literacy instruction, many universities require all preservice teachers enroll in a content area literacy course (Hall, 2005; Herber, 1970); however, research shows that preservice teachers' content area literacy knowledge may not transfer into their instructional practices (Bean, 2000; Hall, 2005). Whereas, elementary teachers receive multiple courses and extensive training in literacy instruction, many secondary programs have one or no course requirements for literacy instruction. By only having one course, pre-service secondary teachers may not have enough opportunities to think about the use of content-area literacy instruction in a thorough nature. Some researchers have argued teacher educators often treat reading in the content areas as a general task, not providing pre-service secondary teachers enough ways to envision literacy instruction in their specific content areas (e.g., Muth, 1993; Wineburg, 2001). Thus, pre-service teachers may need more opportunities to engage in and see models of reading instruction at the secondary level in order to increase their beliefs about reading instruction in the secondary classroom.

Difficulties with current measurement. As shown in Chapter 2, both self-efficacy frameworks (i.e., locus of control and social cognitive theory) struggle with mis-measurement, flawed replications of previous instruments, ceiling effects and a lack of discrimination, and, in some cases, poor statistical methods (detailed in Chapter 2). Additionally, the most commonly used teacher self-efficacy measures (i.e., Teacher

Efficacy Scale [TES], Gibson & Dembo, 1984; Teacher Sense of Efficacy Scale [TSES], Tschannen-Moran & Woolfolk Hoy, 2001) are general in nature and offer little predictive qualities. More specifically, the TES and TSES are unable to provide a causal link between teacher self-efficacy and student achievement as both measures assess teachers' self-efficacy generally which cannot be compared with a specific measure of student achievement. As Parajes (1996, 1997) stated, self-efficacy instruments lose predictability when they measure the construct in a general way. Furthermore, Bandura (1986) articulated that self-efficacy is a context, domain, and task specific construct, and thus should be measured in this manner. Thus, the STERI was designed to be context (i.e., secondary), domain (i.e., four main content areas), task (i.e., content-area reading instruction), and target population (i.e., in-service teachers) specific.

Conclusions

To address the need to assess secondary content area teachers' sense of efficacy for reading instruction, the STERI was designed to capture teachers' beliefs in three specific areas: using content-area reading strategy instruction (Fisher & Frey, 2011; Fisher et al., 2012), motivating students to read (Guthrie & Wigfield, 1999), and teaching struggling readers (Marzano et al., 2001). By designing the STERI focused around these three specific secondary teaching tasks, current measurement issues were not replicated (e.g., lack of specificity, ceiling effects), meaningful group differences were observed, and multiple forms of validity were established.

Specificity in measurement. The creation of a self-efficacy measure focused on secondary reading instruction fills a void present in current research, and offers researchers a way to more specifically measure teachers' sense of efficacy in an effort to link these beliefs to student outcomes. As explained by Bandura (1997) and Pajares (1996), teacher measures must be specific in order to correlate and possibly predict student outcomes; however, current research has been unable to make these connections as the TES and TSES both measure teacher self-efficacy generally, making it impossible to causally link to specific student outcomes. Therefore, the creation of the STERI as a specific measure of secondary reading instruction across the content areas may provide researchers the opportunity to correlate and predict student outcomes from teacher beliefs. Being able to draw a causal path from teachers' sense of efficacy for reading instruction to their students' reading achievement can lead researchers and teacher educators towards creating opportunities for both in-service (e.g., professional development) and pre-service (e.g., literacy courses) to increase: teachers' sense of efficacy, teachers' use of literacy instruction in the secondary classroom, and subsequently their students' reading achievement.

Further, much of the current instrumentation incorrectly aligns with Bandura's (1977, 1986) conception of self-efficacy in relation to social cognitive theory, and are actually more aligned with Rotter's (1966) locus of control theory. This misrepresentation is problematic because Bandura (1977) defined the construct as task-specific and determined it was affected by mastery experiences (1986); whereas, teacher self-efficacy associated with Rotter (1966) focuses more on teacher self-efficacy and student outcome expectancy. Thus, the framework used for scale development influences the types of items written during scale development. As these are two different approaches to defining and measuring the construct of teacher self-efficacy, it becomes important for scale developers to clearly and accurately align their measure and items within one of the two conceptual frameworks.

Ceiling effects. Ceiling effects (i.e., participants answering all items on the higher end of the scale, offering no discrimination in scores) and the inability to discriminate meaningful across participants are additional measurement issues with the current teacher self-efficacy instruments. Ceiling effects lead to many current self-efficacy measures offering little predictive value due to lack of variance because participants tend to score themselves on the higher end of the scale (Klassen et al., 2011). With this over-reporting of high self-efficacy, current measures oftentimes do not show any statistically significant changes over time because teachers' are over-reporting their self-efficacy. This becomes problematic when researchers attempt to demonstrate the effects of professional development on teacher beliefs. As teachers often over-report their self-efficacy on general measures, researchers then have trouble finding statistically significant differences in teacher responses in pre and post studies. This over-reporting leads many researchers to use teacher self-efficacy as a simple descriptive, and often state they were unable to find any meaningful differences over time (Klassen et al., 2011; Wheatley, 2005). Thus creating a situation wherein teacher self-efficacy loses its robustness as a construct and becomes a variable that offers little to no predictive qualities.

Ceiling effects were not present when analyzing responses on the STERI, which may have been due to the specific nature of the items on the measure. Instead of asking teachers to report on their general teaching sense of efficacy, on the STERI teachers were asked to respond to their sense of efficacy for three specific literacy tasks (i.e., Using Content-Area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers). Therefore, the varying difficulty of these tasks and the specific nature of the items may have led teachers to more realistically respond to the items and avoid over-reporting their self-efficacy. More specifically, none of the content areas had mean scores on STERI items that were high enough to suggest ceiling effects in scoring. This may provide support that more specifically written items can aid in reducing ceiling effects and lead to more substantive findings when examining teacher self-efficacy.

Verbal anchor scores. Moreover, the STERI used a six-point verbal labeling in lieu of numerical anchor scores to avoid participants inferring a meaning associated with numerical values (e.g., Tourangeau et al., 2000). One of the benefits of using a six-point scale is the lack of a true middle; forcing participants to select a higher or lower sense of efficacy for each item. In addition, by using verbal labels participants are less likely to lump responses together (e.g., viewing 1-3 as all representing “not at all” when only one was meant to reflect this response) and provide a single response for each item. The lack of ceiling effects and the variability in responses on the STERI may have been attributed to the verbal anchor scores.

Using verbal anchors provides participants with a term associated with each anchor (i.e., *Not at all; Very little; Somewhat; Moderately; Quite a bit; and Completely*) instead of the more widely adopted use of numerical anchor scores where only certain anchors have terms associated with them (e.g., 1 *strongly disagree*, 2, 3 *neither disagree* *or agree*, 4, 5 *strongly agree*). When scales only use numerical anchor scores and do not provide terms associated with each word, the anchors become more arbitrary because participants are tasked with inferring what a two or four means without having an actual term associated with the scale (e.g., Tourangeau et al., 2000). Therefore, the variability on the STERI may have been due in some ways to the fact that teachers were able to associate a term instead of a numerical value to each item within the STERI scale, leading to a lack of ceiling effects and greater variability in responses.

Meaningful group differences. Using one-way ANOVAs for each of the three STERI subscales demonstrated that the STERI was able to discriminate across content areas (see Chapter 4 for a detailed explanation). In particular, each of the three subscales followed the same pattern in regards to teachers' mean difference in scores: English, social studies, science, and mathematics; therefore English teachers had the highest sense of efficacy in reading instruction, and mathematics teachers had the lowest sense of efficacy in these areas. These differences were theoretically meaningful as they followed the hypothesis that this pattern would emerge based on the amount of reading done within each of the four core content areas. As the amount of reading differs based on content area, it was assumed that teachers' level of self-efficacy for reading instruction would correlate with the amount of reading necessary in that content area. In particular, the English content area has a large amount of reading required, thus it was assumed these teachers would have higher self-efficacy for reading instruction than other content area teachers. The amount of reading done in each content area follows a pattern of: English, social studies, science, and mathematics; therefore, teachers' level of self-efficacy for literacy instruction was hypothesized to mirror this pattern.

Moreover, these conclusions may have arisen due to the specific task measurement of the STERI; as these theoretically meaningful patterns did not arise on the TSES. Further, although the STERI used a 6-point Likert scale and the TSES used a 9-point Likert scale, the STERI had greater mean differences and larger effect sizes (d = .48 – 2.26) than the TSES (d = .40-71). Suggesting that measuring teacher self-efficacy more specifically may lead to more statistically and practically significant differences across groups within the sample. By assessing teacher self-efficacy for reading instruction at a specific level, theoretical hypotheses were confirmed; thus, providing further support that specificity in measurement is of paramount importance for effective self-efficacy measurement (Bandura, 1997; Pajares, 1996),

Group differences were also found when examining responses on the STERI based on number of literacy professional development sessions attended and number of literacy courses taken for credit (see Chapter 4 for details). For both variables, statistically significant group differences were found across subscales for teachers with zero professional development/literacy courses and their colleagues with one to five sessions/courses. These findings align with the literature that states that teacher self-efficacy for literacy instruction at the secondary level can be influenced by professional development (e.g., Fisher et al., 2002) and that literacy requirements should be enhanced for secondary teachers in order to increase positive beliefs regarding literacy instruction (e.g., Shanahan & Shanahan, 2008).

By finding meaningful group differences, the STERI provides evidence that writing items based on literature and not from previous measures can create items that discriminate in a theoretically meaningful way; therefore, more accurate hypotheses can be drawn about how items may discriminate. This was demonstrated in this study when examining these variables (i.e., number of literacy professional development sessions, number of literacy courses taken) and the TSES. No meaningful group differences were found using the TSES, providing further evidence that when teacher self-efficacy is measured more specifically the possibility to find meaningful group differences may occur. This contributes to the teacher self-efficacy research by providing a possible solution to the issue of participants over-reporting of their self-efficacy and the lack of meaningful differences amongst participants (e.g., Klassen et al., 2011; Wheatley, 2005) by creating task-specific items that are more closely aligned with Bandura's (1977) original conception of self-efficacy.

Summary. In sum, the STERI was able to show meaningful and theoretical sound group differences, whereas the TSES could not; thus, the STERI was measuring teachers' sense of efficacy differently than the TSES. By assessing teacher self-efficacy for specific teaching tasks, the STERI discriminated across content areas, in a theoretically meaningful way; however, the TSES measured general teaching self-efficacy and therefore did not discriminate in a meaningful way. Also, the STERI was able to demonstrate that teacher's with at least one to five literacy professional developments or literacy courses had a statistically significantly higher means across all three subscales than teachers with neither of these respective opportunities. These findings mirror the literature that demonstrates the need for professional development experiences and literacy courses to enhance secondary content area teachers' beliefs about literacy instruction in the classroom (Biancarosa & Snow, 2006; Fisher et al., 2002; Olson & Truxaw, 2009; Shanahan & Shanahan, 2008).

Professional development. By approaching professional development as a preparedness variable I was able to align my results with other researchers who found professional development contributes to teachers' sense of efficacy. More specifically, these findings align with the work of Tschannen-Moran and Johnson (2011) who found that literacy professional development experiences contributed to teachers' self-efficacy. By replicating these findings, the notion that professional development experiences can be designed to foster teachers' self-efficacy is strengthened.

Further, findings of group differences on the STERI demonstrate that teachers' mean score on the STERI were statistically significantly higher when teachers attended one to five sessions, and this differences was a full point when they attended six to eleven professional development experiences compared to those who attended zero. Hence, teachers who had attended literacy professional development sessions had higher mean scores on the STERI than those teachers who had not attended any sessions. These findings suggest that teachers do not need to be overwhelmed with double digit literacy professional development sessions in order to see growth in self-efficacy beliefs. Rather, these experiences should be more targeted for the areas in which teachers need the most enhancement of beliefs.

The creation of the STERI aids researchers in creating and assessing targeted professional development focused on improving teachers' use of content-area literacy instruction in the secondary classroom. By measuring teachers' self-efficacy beliefs more specifically, researchers can ascertain these beliefs and target professional development opportunities towards specific literacy skills. Furthermore, researchers can also gauge the effectiveness of professional development by examining changes in teachers' sense of efficacy using the STERI both before and after professional development sessions as a diagnostic tool.

Pre-service literacy courses. Previous studies have not shown a significant contribution of level of education and preparation in predicting self-efficacy (i.e., Raudenbush et al., 1992). However, when studies have examined teachers' level of preparedness (i.e., "I feel well prepared to teach this class"), the results more consistently demonstrate a statistically significant relation to teacher self-efficacy (see Raudenbush et al., 1992 and Ross et al., 1996 for examples). Therefore, by approaching preparedness as the number of literacy courses taken for credit I was able to find support for past research as well as discover meaningful group differences based on teacher preparedness.

Teachers with no literacy courses scored statistically significantly lower than all other groups on the STERI. This finding suggests that with even one literacy course secondary teachers feel more self-efficacious at using literacy instruction in the secondary classroom. Moreover, following a similar pattern as with professional development, teachers with six to eleven courses had a mean a point and a half higher than those teachers with no courses. This suggests that pre-service teachers do not need to take a multitude of literacy courses in order to increase their self-efficacy beliefs, but rather they need literacy instruction in at least one course.

With respect to pre-service teachers, if teacher education programs increase their emphasis on literacy instruction in secondary classrooms, the STERI may be used as a diagnostic tool to assess the effectiveness of the redesigned courses. Therefore, the STERI offers researchers and teacher educators a means by which they can more specifically measure teachers' beliefs about their own effectiveness in literacy instruction at the secondary level. Further, teacher education programs can use the STERI to determine if one stand-alone course is enough to enhance teacher beliefs regarding their effectiveness at literacy instruction.

Validity. Many of the current teacher self-efficacy measures suffer from using a mixed sampling of teachers to establish validity (e.g., Tschannen-Moran & Johnson, 2011) and an overall lack of convergent (e.g., Paneque & Barbetta, 2006) and discriminant validity (e.g., Haverback & Parault, 2011). To address these concerns, I followed Gehlbach and Brinkworth's (2011) six-step approach to frontload validity. Further, by creating the STERI based on theory I was able to avoid many of the psychometric issues confounding current instrumentation. Finally, by correlating the STERI with both convergent (i.e., TSES and Goal Orientations for Teaching [GOT]) and discriminant (i.e., Invitations to Self/Others) scales, I was able to show that the STERI was measuring teacher self-efficacy in a similar but also different way than the established TSES and GOT scales. More importantly, the STERI was one of the few self-efficacy measures that provided clear discrimination across participants. This provides further support for the creation of future teacher self-efficacy instruments that are designed to assess the construct in context, domain, task, and target population specific ways (Bandura, 1997; Pajares, 1996).

Face and content validity. Using a literature review, teacher interviews, expert review, and cognitive interviews, I was able to establish face and content validity before conducting exploratory analyses (Gehlbach & Brinkworth, 2011). By frontloading validity I was able to qualitatively revise and refine the STERI through multiple iterations of expert review and cognitive interviews. This helped me to not have the same qualitative issues as some of the previously developed self-efficacy measures (e.g., TES; Gibson & Dembo, 1984). By grounding the items of the STERI within the adolescent literacy literature (e.g., Biancarosa & Snow, 2006; Guthrie & Wigfield, 2000; Vacca, 2002), I was able to create new items that were not based upon previous instruments. Creating items in this way ensured that I did not replicate past issues, and also helped in developing hypotheses regarding item discrimination grounded in the literature. Additionally, by using teacher interviews I was able to confirm what I found in the literature with what practicing teachers were experiencing in the classroom. Creating items using the literature and teacher interviews provides researchers the opportunity to create more qualitatively sound instruments grounded in theory.

By conducting two rounds of expert review (n = 10), I was able to establish content validity for the STERI and further revise the items before quantitative analyses occurred. Content validity helps researchers determine if the newly formed instrument is measuring the intended construct (Crocker & Algina, 1986; DeVellis, 2003). By establishing content validity I was able to determine if the STERI was in fact measuring both literacy instruction and teacher self-efficacy in a theoretically meaningful way. Further, by having two rounds of expert review I was able to make multiple rounds of qualitative revisions to items; ensuring that the most concise and well written items were included on the final version of the STERI. Using multiple rounds of expert review provides scale developers numerous opportunities to get qualitative feedback in order to prepare the most theoretically meaningful items.

Another means by which I established validity was through two rounds of cognitive interviews (n = 12). Cognitive interviews serve as a way for researchers to determine how participants from the target population interpret, understand, and respond to items on a scale (Campanelli, 1997; DeMaio & Rothgeb, 1996; Willis, 2005). The STERI was the first teacher self-efficacy measure that used cognitive interviews as a way to establish face validity. By engaging high school teachers from each of the four core content areas (i.e., English, science, social studies, and mathematics) in cognitive interviews, I was able to determine how teachers understood and responded to items on the STERI. These interviews provided insights into how practicing teachers viewed the items on the scale and created the opportunity for me to ask participants questions as they responded to items.

This level of targeted and specific interviewing gave me insights into how teachers would think about and respond to the items on the STERI. Also, I was able to physically see and hear when participants found items difficult and ask them probing questions to determine what about the wording of items made them difficult to understand. Conducting these interviews helped me to revise all items, establish face validity for the STERI, and see firsthand how my instrument was understood by the target population (Willis, 2005). This was a key step in the development of the STERI as it created opportunities to engage with the target population and ensure the best version of the items were included before conducting a large-scale quantitative analysis (Gehlbach & Brinkworth, 2011).

Summary. Adapting Gehlbach and Brinkworth's (2011) six-step approach to scale development (see Chapter 3 for details) allowed me to successfully frontload validity for the STERI. By using the literature and teacher interviews, I was able to construct new theoretically meaningful items. Expert reviews helped me to further revise these items and ensure they were correctly measuring my intended construct. Conducting cognitive interviews enabled me to see first-hand how my target population responded to items on the STERI. Future researchers should consider modeling their scale development off of this framework (Gehlbach & Brinkworth, 2011) as it frontloaded validity on the STERI and aided in creating a qualitatively sound instrument.

Target population. As hypothesized by Bandura (1986) the experiences a person has firsthand (i.e., mastery experiences) may influence one's self-efficacy more so than any other persuasion. Therefore, when establishing validity it is important that the sample represents participants who all have similar mastery experiences. In relation to teaching, it is important that validation studies only occur within one specific group (i.e., pre-service or in-service teachers). The need for this separation based on mastery experiences, comes from the literature that states pre-service teachers often over-estimate their teaching self-efficacy as they have very little mastery experiences in teaching (e.g., Henson, 2002). However, many of the current instruments use a mixed sampling of pre-service and in-service teachers to establish validity (e.g., TES, Gibson & Dembo, 1984; TSES, Tschannen-Moran & Woolfolk Hoy, 2001). Not only is this problematic because these teachers have different mastery experiences, but research has shown that the factor structure of the TSES is not consistent across pre-service and in-service teachers (Duffin et al., 2012; Fives & Buehl, 2010; Henson, 2002).

My sample consisted of only in-service teachers to eliminate the possibility of mixed results due to a sampling of pre-service and in-service teachers. Further, to ensure that the sample was as homogenous as possible based on mastery experiences, only high school (i.e., grade 9-12) teachers were used in the initial validation study. By limiting the sample to only high school practicing teachers, I was able to ensure that all teachers had similar mastery experiences and I limited the extraneous variables that could have confounded my results. Thus, future studies should validate instruments using only samples that have the same mastery experiences (Bandura, 1986) in order to alleviate measurement issues. Moreover, scale developers should attempt to validate instruments using as homogeneous a sample as possible to avoid finding confounding results.

Convergent validity. Convergent validity is used to ensure that the newly established instrument is measuring a similar construct as an already validated instrument (Crocker & Algina, 1986). Although scale development research emphasizes the establishment of convergent validity as a necessary step (e.g. DeVellis, 2003), some teacher self-efficacy measures do not adhere to these suggestions (see Table 3).

However, convergent validity is an essential component of scale development and should be used by all researchers who are seeking to develop measures.

To establish convergent validity, correlational analyses with the STERI subscales and the TSES and the GOT were conducted. The moderate correlations between the STERI and the TSES demonstrated that the STERI was related, but not too similar, to the TSES; thus, they are measuring similar constructs, but not too similarly. Further, when attempting to establish discriminant validity between the STERI and the Invitations to Self/Others subscale, convergent validity was actually established as the STERI subscales had a similar relationship to the Invitations subscale as the TSES subscales. Because the TSES was an already validated measure of teacher self-efficacy, seeing similar patterns in relationships in the STERI and TSES subscales and their relation to the Invitations Subscales provided further evidence that the STERI was measuring a similar construct as the TSES.

The GOT mastery and relational goals subscales were also used to provide evidence of convergent validity. Much like the TSES, relationships between the STERI and the GOT subscales were as expected with no correlations exceeding $r = .50$; confirming the relationship between teacher self-efficacy and goal orientations for teaching, but without too strong a relationship to suggest an overlap in measurement. Yet, very similarly to what was stated above, the TSES and the STERI both had similar relationships to the GOT subscales. This established further convergent validity for the subscales of the STERI measuring a similar construct as the TSES.

Summary. The STERI is one of the only teacher self-efficacy measures that sought to establish convergent validity using both a teacher self-efficacy scale and another educational construct (i.e., GOT, Butler & Shibaz, 2014). By using both scales, I was able to establish that the STERI was measuring teacher self-efficacy, and also that goal orientations and teacher self-efficacy are weak to moderately positively correlated. These findings contribute to the research on teacher self-efficacy because it offers another construct that has a relationship with teachers' sense of efficacy, and can provide opportunities for future research to attempt to establish more thorough profiles of teachers' beliefs and their influences on classroom practices.

Discriminant validity. Many of the current measures of teacher self-efficacy never established discriminant validity which offers a way for researchers to demonstrate their instrument is not measuring something it is not intended (DeVellis, 2003). As many studies did not establish discriminant validity, one issue that arose during this study was the lack of measures to use to establish discriminant validity; an issue I will discuss in the limitations section below. In my attempt to establish discriminant validity, the STERI subscales were correlated with the Invitations to Self/Others. Although I only found discriminate validity for the STERI reading strategies factors (i.e., a weak non-significant correlation; r = .11) the patterns seen between the STERI subscales and the TSES subscales in relation to the Invitations to Others/Self subscales provided further convergent validity, as both subscales correlated similarly to the Invitations to Self/Others subscales.

Additionally, discrimination from the TSES was also confirmed when examining group differences on (i.e., content area, number of literacy professional development sessions, number of literacy courses taken) on both the STERI and the TSES. Although the STERI was able to discriminate across groups in a theoretically meaningful pattern, the TSES was not. Providing further evidence that the STERI and the TSES were measuring teacher self-efficacy differently. In addition, the STERI data showed discrimination across content areas in a similar pattern and the TSESs inability to do so, established that the STERI measured teacher self-efficacy in a different way than the TSES; providing discriminant validity for the STERI.

Summary. By frontloading validity and following the suggestions of scale development researchers (e.g., Crocker & Algina, 1986; DeVellis, 2003; Gehlbach & Brinkworth, 2011), I was able to create the STERI to avoid many of the issues that have plagued past teacher self-efficacy measures. Additionally, establishing convergent validity using both a teacher self-efficacy scale and another construct offers researcher the opportunity to develop more thorough profiles of teachers. Future scale developers may want to follow the steps used to develop and validate the STERI as these led to having multiple forms of validity (i.e., face and construct) established before exploratory analyses occurred to limit any qualitative issues with items.

Limitations

Although I conducted this study using rigorous qualitative and quantitative methods, limitations still arose during the development and validation of the STERI. One limitation of the current study was that the sample had an overrepresentation of English teachers (e.g., 40% of the sample were English teachers) and did not align with the 2013 School and Staffing Survey (SASS). Although the SASS lists English teachers as the highest percentage of high school teachers, the percentages across content areas on the SASS are more closely aligned than that of the STERI participants. As a former English teacher, many of the initial teachers with whom I asked to participate in the study were high school English teachers. This led to a heightened amount of English teachers in the study. Chi-square analysis revealed that most of the English teachers in the study were female which led to a heightened amount of females in the study. Therefore, future research should seek to more closely mirror the SASS in regards to the percentage of teachers within each content area participating in the study to have samples that more closely align with the population.

Another limitation was the lack of discriminant validity established by the Invitations to Self/Others. The lack of clear discrimination between the STERI and the Invitations subscales was problematic because it did not clearly establish the theoretically difference between the constructs of self-efficacy and invitations to self/others. By not establishing discriminant validity, I was unable to clearly establish that the STERI was not measuring the construct of Invitations to Self/Others. When examining the items within the Invitations to Self/Others scale it became evident that the moderate relationship between the STERI and the Invitations subscales should not have been unexpected.

Although studies have found that invitations are distinctly different from the social persuasions (i.e., one of the four hypothesized sources of self-efficacy; Bandura,

1997) because they are messages sent rather than received (Morris, 2010; Usher & Pajares, 2006); I was unable to find this distinction in measurement. Upon review of the items and when thinking about the messages a teachers send to themselves/others, it makes sense that teachers' messages could be related to self-efficacy. For example, if as a teacher I feel confident in my abilities in teaching, than I would send myself more positive messages about my own teaching. Hence, when reviewing the scale and thinking through the relationships that were demonstrated between the STERI and the Invitations to Self/Others, it became evident the Invitations to Self/Others was not the best scale to establish discriminant validity.

This becomes problematic because proper scale development seeks to establish four types of validity: face, content, convergent, and discriminant to ensure that the newly created instrument is measuring only what it is intended to measure (Crocker & Algina, 1986; DeVellis, 2003; Gehlbach & Brinkworth, 2011). Therefore, discriminant validity must be established to ensure that the STERI is not measuring another construct as well as teacher self-efficacy for reading instruction. These issues arose primarily due to the fact that no other measures had been created that were as specific as the STERI, none centered around adolescent literacy instruction, and most teacher self-efficacy scale development literature did not establish discriminant validity. Additionally, when examining the teacher self-efficacy scale development literature, many of the scales used to establish discriminant validity were not related to education (e.g., the TSES used the Work Alienation Survey; Tschannen-Moran & Woolfolk Hoy, 2001). Thus, it was difficult to find a measure that was related to the education field but did not overlap with self-efficacy. Future research must be conducted to find a better scale within the field of education to establish discriminant validity from the STERI.

Finally, the order of the survey instrument may have caused the TSES Instructional Strategies factor to function similarly to the STERI Using Content-Area Reading Strategy Instruction factor, as the TSES items came after the STERI items. As both sets of items used the term *strategies* teachers' may have responded to both sets of items similarly; even though the STERI items specify *reading strategies* and the TSES uses *instructional strategies.* By having the STERI items come before the TSES items, teachers' may have responded to the TSES items thinking about reading strategies as they had already responded to those items on the STERI; thus, causing the two subscales (i.e., STERI, Using Content-Area Reading Strategy Instruction and TSES, Instructional Strategies) to be superficially highly correlated. Therefore, in future studies the items on the survey package need to be randomized to determine if the order of the items affects the nature of participant responses; particularly to the STERI Using Content-area Reading Strategy Instruction factor and the TSES Instructional Strategies factor. More specifically, if the TSES factor comes first will teachers respond that the two subscales are correlated and will their responses to the items differ because they will not have previously read any reading strategy items.

Implications for Research

As the need to causally link teachers' self-efficacy and student achievement becomes paramount, the creation of the STERI provides the first form of measurement that is based on Bandura's (1977, 1986) conception of self-efficacy and is context, domain, task, and target population specific. Further, the creation of the STERI offers researchers a means by which to: develop literacy professional development targeted around the three main factors of the STERI (i.e., Using Content-Area Literacy Instruction, Motivating Students to Read, Teaching Struggling Readers), examine effectiveness of said literacy professional development, and determine if there is a causal link between teacher self-efficacy, teacher instruction, and student achievement all of which were limited before the development of the STERI.

Previous research (e.g., Fisher & Ivey, 2005) has shown that teachers with little confidence in their ability to integrate reading instruction in their content area will avoid this type of instruction. However, further research should be conducted using the STERI to determine if teachers' responses on the STERI correspond to their reading practices within the classroom. This link can be established by first providing teachers the opportunity to take the STERI and then examining their reading practices through classroom observations. Qualitative researchers have denoted that classroom observations can be an accurate way to examine the instruction in which teachers are engaging in their classroom (Evertson & Green, 1986; Maxwell, 2012; Patton, 2002). Therefore, using classroom observations, researchers can determine if there is a link between teachers' responses on the STERI and the level of reading instruction that occurs in their classroom. Moreover, this research adds to the literature that teachers' sense of efficacy influences their instruction (Guo et al., 2010; Tschannen-Moran et al., 1998), but in a manner that targets both self-efficacy beliefs and instruction in a more specific manner.

As the first literacy specific measure that was created using Bandura's (1977) conception of self-efficacy in a context, domain, and task specific way the STERI provides a more precise level of measurement than all current teacher self-efficacy measures. This difference was seen when examining group differences and ceiling effects with both the STERI and the TSES. When examining group differences the STERI, patterns followed a theoretical pattern and no ceiling effects occurred; however, no meaningful patterns arose when examining group differences on the TSES, and ceiling effects were present. Therefore, when teachers were given very specific teaching tasks (e.g., In your content area how well can you model effective reading strategy use?) on the STERI, they did not select anchors towards the higher end of the scale; thus, negating ceiling effects. This provides support for Bandura's (2006) guide for constructing self-efficacy scales, which articulated the need to write items to represent tasks of varying degrees of difficulty to reduce the risks of ceiling effects and provide variability in responses. Across all four content areas, ceiling effects were not present and mean group differences demonstrated that variability of responses occurred. Thus, future researchers should seek to write items that assess specific teaching tasks of varying degrees of difficulty in an attempt to decrease ceiling effects.

Additionally, the STERI provides researchers and teacher educators a means by which to design and assess the effectiveness of targeted literacy professional development. More specifically, if researchers can use the STERI as a diagnostic tool, they can determine which area (i.e., Using Content-Area Literacy Instruction, Motivating Students to Read, Teaching Struggling Readers) teachers need the greatest improvement in their self-efficacy beliefs, and can design professional development meant to focus on these key areas. Not only can researchers design these targeted professional development experiences, but they can also determine the effectiveness of these experiences at increasing teachers' sense of efficacy in literacy instruction using a pre and post test.

Further, with the creation of the STERI, researchers can attempt to draw a causal link between teachers' sense of efficacy for reading instruction in the secondary classroom, reading instruction, and their students' reading achievement. More specifically, as the STERI measures teachers' sense of efficacy specifically, researchers now have the opportunity to attempt to draw causality between teachers' sense of efficacy for reading instruction, their use of reading instruction, and their students' reading achievement. Further, if this link can be established than researchers can attempt to design professional development experiences for both pre-service and in-service teachers meant to increase teachers' self-efficacy in literacy instruction in an attempt to increase the use of literacy instruction in the secondary classroom and subsequently increase their students' reading achievement.

Implications for Practice

Researchers have shown that sustained professional development, where secondary teachers learn and practice content area literacy instruction, has led to positive beliefs about literacy instruction being included in the secondary classroom, increased use of literacy instruction in classrooms, and increased student achievement (Biancarosa & Snow, 2006; Fisher, 2001; Heller & Greenleaf, 2007). This supports the idea that beliefs can be changed through systemic and ongoing professional development and support. As Pajares (1992) suggested, when confronted with new knowledge, teachers process and interpret that knowledge and decide whether it aligns with their beliefs. Therefore, teachers must be given the opportunity to not only engage with knowledge about literacy instruction in the content areas, but have the space to address their previously held beliefs about literacy instruction and their personal sense of efficacy in literacy instruction.

Fisher and Ivey (2005) stated that researchers and educators must change teachers' experiences with reading, writing, and learning if they want to change content area teacher's perceptions about their role in literacy instruction. In order to change beliefs, teachers must be given the opportunities to examine previously held beliefs and work towards new conceptualizations. Vacca (2002) echoed this sentiment, "In addition, schools need to implement ongoing staff development efforts, including instructional strategy workshops, self-study, teacher inquiry projects, and action research in the various content areas in the middle and high school curriculums" (p. 10). Fisher and Ivey (2005) articulated this need and discussed various ways by which both preservice (e.g., increased support and modeling of reading strategy use in content classes within methods courses) and in-service teachers (e.g., professional development for teachers to discuss and see models of strategy use in content classes) can explore, practice, and engage in reading strategy instruction within their content area classrooms. Understanding teachers' beliefs regarding literacy instruction and their self-efficacy in teaching reading strategies provide a means to more effectively cater professional development to teachers' needs.

With the insurgence of the Common Core State Standards (CCSS) and the push for literacy instruction in the secondary classroom to address poor adolescent reaching achievement, teacher educators must construct a curriculum that infuses literacy and content instruction for preservice and in-service teachers. Therefore, preservice secondary teacher education programs should attempt to incorporate more comprehensive literacy course requirements and opportunities for preservice teacher to engage in literacy strategy instruction across content areas. Additionally, schools need to implement sustained professional development for in-service teachers to gain knowledge, observe models, and practice the use of literacy and content instruction in tandem. Both preservice teacher education programs and in-service sustained professional development focused on literacy instruction in secondary education can provide teachers the knowledge, practical skills, guidance, and ability to change or enhance teacher beliefs about literacy instruction in their classrooms.

As teachers, both preservice and in-service, confront this new knowledge the opportunity arises for teacher educators to measure teachers' beliefs about literacy instruction in the content areas and their personal beliefs about their ability to effectively execute literacy instruction. Thus, the need arises for researchers to develop instruments to ascertain these specific contextual beliefs. These measures can be used to determine if pre-service teacher education and professional development programs are effective in fostering teachers' personal sense of efficacy in using reading instruction.

Recommendations for Future Research

This study focused on the development and validation of the STERI and leads to future research to confirm these findings so that the STERI may be used with multiple samples. Future research needs to be conducted to further confirm the three factor structure of the STERI (i.e., self-efficacy for: Using Content-area Reading Strategy Instruction, Motivating Students to Read, and Teaching Struggling Readers) using a new sample of high school teachers. As the first stage of factor analysis (i.e., Exploratory Factor Analysis, EFA) has already been conducted, future studies with high school teachers should use the second stage of factor analysis (i.e., Confirmatory Factor Analysis, CFA). CFA tests specific hypotheses the researcher has about the numerical values of the parameters (Crocker & Algina, 1986). Therefore, a future study using CFA would be conducted using the three-factor structure found in the initial EFA.

Moreover, a more theoretically distinct scale should be found to provide further discriminant validity for the STERI. One possible solution regarding discriminant validity may be to conduct another review of the literature to determine if domain-specific measures of teacher self-efficacy (outside of literacy) based on social cognitive theory have been created. At the time of this study, no other domain-specific measures had been created based off of social cognitive theory. If newly created domain-specific measures have been created, validation can occur by examining correlations and responses on the STERI and the newly created instrument to attempt to more firmly establish discriminant validity. This discrimination should occur because according to Bandura (1977, 1986), self-efficacy is task-specific. Therefore, if another scale was developed in a different domain that measured specific teaching tasks within that domain, teacher self-efficacy should be different on the STERI and the newly created measure (i.e., teachers' sense of efficacy would be unrelated on the two different teaching tasks).

Another solution could be the use of the work alienation survey (Forsyth & Hoy, 1978) that was used to establish discriminant validity for the TSES (Tschannen-Moran & Woolfolk Hoy, 2001). However, one explanation for the lack of constructs to establish discriminant validity regarding teaching self-efficacy based on Bandura's (1977) may be the notion that self-efficacy is a construct that influences humans through triadic reciprocation and therefore all educational constructs will overlap with teaching self-efficacy in some form or another. Thus, it may become necessary to find a scale that measures a construct outside of the educational field to firmly establish discriminant validity.

After a CFA is conducted and discriminant validity further established, the STERI can be utilized to assess high school secondary teachers' sense of efficacy in reading instruction across the content areas. Moreover, group differences must be examined (i.e., content area, number of literacy professional development sessions attended, number of literacy courses taken) with a new sample to determine if group differences in this sample were a by-product of sampling, or true discrimination by the STERI. Additionally, the significant teacher and environment variables in this study (i.e., gender, level of education, grade level taught, school Title I status, percent of below-grade readers) need to be examined with a new sample of teachers to determine if these variables are statistically significant with different samples or only this particular sample; especially as these findings both confirmed (e.g., Klassen & Chiu, 2010) and refuted (e.g., Tschannen-Moran & Hoy, 2007) previous findings (see Chapter 4 for details).

Additionally, the STERI factor structure needs to be confirmed using a sample of middle school teachers. As the STERI was intended for use with all secondary teachers, a new validation study should occur with middle school teachers to ensure the factor structure, reliability, and validity of the STERI hold up when used with middle school teachers. However, knowing the great variability that exists when discussing middle school, for all validation studies, samples should be as homogeneous as possible. Thus, specific criterion for participant inclusion into the studies must be adhered to by the researchers. This research should initially use a sample of middle school teachers to conduct an EFA to explore how middle school teachers respond to the items on the STERI. Upon completion of an EFA, a new sample of middle school teachers should be recruited to conduct a CFA and further confirm the factor structure of the STERI.

Further, group differences should also be examined with middle school teachers to determine if the same patterns seen with high school teachers are present with middle school teachers. Moreover, as some middle school teachers work in cross content area *teams*, group differences should also be examined based on teachers who work in teams and those that do not. These findings may provide some insight into if working in teams leads to higher levels of self-efficacy for reading instruction regardless of teachers' primary content area of instruction. If teachers who work in teams score statistically significantly higher on the STERI than teachers who do not, the argument could be made that high school teachers' sense of efficacy for literacy instruction may increase if they work in cross content area teams.

Once the factor structure of the STERI has been confirmed and discriminant validity more firmly established, the STERI can be used as a reliable and valid instrument in future studies. As stated previously in this chapter, the STERI can then be used as a diagnostic tool to design and assess the effectiveness of literacy professional development and pre-service literacy courses. Additionally, the creation of the STERI offers researchers the opportunity to attempt to find a causal link between teacher self-efficacy, teacher instruction, and student achievement.

Therefore, future studies should use the STERI to assess teachers' sense of efficacy for reading instruction, observation data to examine teacher instruction (Evertson & Green, 1986; Maxwell, 2012; Patton, 2002), and student reading achievement data to assess if teacher self-efficacy influences reading instruction. These studies would add to the literature that demonstrates that reading instruction influences student reading achievement (Fisher et al., 2012; Nokes & Dole, 2004; Underwood & Pearson, 2004). If these links can be established, the implications of these findings go beyond the field of literacy; suggesting the need for all domains to create more specific teacher self-efficacy measures to examine if this causal link is present. If teacher self-efficacy is shown as a predictor for teacher instruction and student achievement, than the field of teacher self-efficacy can expand as researchers and teacher educators seek to enhance these beliefs to impact instruction and improve student achievement.

Summary

The STERI is the first teacher self-efficacy instrument that was designed to assess Bandura's (1977, 1986) conception of self-efficacy and was context, domain, task, and target population specific. Additionally, it is the first teacher self-efficacy measure designed specifically for secondary content area teachers within the domain of literacy. Following Gehlbach and Brinkworth's (2011) model and using suggestions from scale developers (e.g., DeVellis, 2003), the STERI was developed without replicating past measurement issues and provides a model for future scale developers. Qualitative analyses occurred using: literature review, teacher interviews, synthesis of this information, item writing, expert review, and cognitive interviewing. Engaging in these qualitative analyses before conducting an EFA created the opportunity to frontload validity and ensure that the most theoretically and qualitatively sound items were included on the STERI before quantitative analyses were conducted. Future studies should seek to replicate these steps when creating scales in order to create more qualitatively sound and valid instruments.

With the creation of the STERI, professional development can be targeted for both pre-service and in-service teachers to develop the particular areas of reading instruction (i.e., using content-area reading strategy instruction, motivating students to read, and teaching struggling readers) needed. Using the STERI as a diagnostic tools provides researcher and teacher educators a means by which to design and assess the effectiveness of professional development for both pre-service and in-service teachers. More specifically, the STERI can assess secondary teachers' specific self-efficacy beliefs in an attempt to: increase teachers self-efficacy in reading instruction, enhance the use of reading instruction at the secondary level across the content areas, and subsequently increase secondary students’ reading achievement.

Appendix A

George Mason University IRB Exempt Letter

GEORGE MASON UNIVERSITY

Office of Research Integrity and Assurance

Research Hall, 4400 University Drive, MS 6D2, Fairfax, Virginia 22030
Phone: 703-993-5445; Fax: 703-993-9590

DATE: November 12, 2014

TO: Michelle Buehl

FROM: George Mason University IRB

Project Title: [682512-1] Development and Validation of a Teacher Sense of Efficacy Scale in Reading Instruction

SUBMISSION TYPE: New Project

ACTION: DETERMINATION OF EXEMPT STATUS

DECISION DATE: November 12, 2014

REVIEW CATEGORY: Exemption category #2

Thank you for your submission of New Project materials for this project. The Office of Research Integrity & Assurance (ORIA) has determined this project is EXEMPT FROM IRB REVIEW according to federal regulations.

Please remember that all research must be conducted as described in the submitted materials.

Please note that any revision to previously approved materials must be submitted to the ORIA prior to initiation. Please use the appropriate revision forms for this procedure.

If you have any questions, please contact Karen Motsinger at 703-993-4208 or kmotsing@gmu.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within George Mason University IRB's records.

Appendix B

Semi-Structured Interview Questions for Social Studies Teachers

1. What do you think is the best way to help students read to learn in social studies?
2. What knowledge do you think teachers need in order to incorporate reading comprehension in social studies?
3. [In response to what a teacher said to question 2] And what is it that you need? Like what do you need to know?
4. So then, what is it, you talked about needing to be able to meet the needs of students, what does that require?
5. Do you think having a strategy to link onto helps?
6. So how confident or comfortable do you feel in teaching different strategies?
7. Are there benefits to incorporating literacy instruction in social studies if there are any?
8. What do you think is the best way for students to read to learn?
9. Do you think that modeling the strategy for them [students] helped?
10. What kind of supports do you think future social studies teachers need to incorporate literacy instruction in social studies?
11. Do you think you need, like did you ever get anything, on diagnosing why they [students] can't read like when a kid can't tell you what's going on?
12. And is it hard for you to do reading, literacy instruction in social studies?
13. [In response to question 12]: Teacher, "It's just different. Um, it's necessary." Interviewer: And why do you think it's necessary?

Appendix C

Table 5

Quotes from teacher interviews that support the inclusion of the three factors

Struggling Readers	Strategy Use	Motivation to Read
"I think you have to use a variety of modalities to get these kids to learn because in our classes we really have a mixture of students who are at different reading levels. It would be nice to have everyone at the same reading levels but that is a dream, not a reality." Teacher 1	"And so we have to use different strategies, a couple of different strategies to try to reach everybody." Teacher 1	"The teachers need to know that because you don't want the kid who knows it to get bored, you want to teach them but you don't want the kid who struggles to understand to have trouble learning and be overwhelmed by something and lose faith in the system and lose hope." Teacher 1
"They [teachers] need to understand there are kids who are going to understand really really basic English, there are kids who do not understand English, and there are kids who understand English at a 11th and 12th grade reading level." Teacher 1	"Well obviously they need a knowledge of the different modalities and different strategies, because obviously these kids need different things because they are at different reading levels." Teacher 1	"Motivation is the toughest thing. That aspect of education is really really hard. We have to tell kids why history is important to learn and some kids will know that and will care and so throughout the year, we have to go back periodically. This is the connection to things in the modern world and this is how this will benefit you." Teacher 1

"The teachers need to know that because you don't want the kid who knows it to get bored, you want to teach them but you don't want the kid who struggles to understand to have trouble learning and be overwhelmed by something and lose faith in the system and lose hope." Teacher 1	"You have to understand that connection because otherwise they won't understand the relevance of having to take history." Teacher 1	In response to liking literature circles: "Why do I like, because I didn't think they were going to work and then they worked and I was like oh my god, so I kinda want to try it again at some point. They seemed engaged." Teacher 2
"Or even if you understand what you're reading, if you can't put it in the context of what they're trying to say then it doesn't quiet work for ya." Teacher 2	"The strategies we have used them before and we are certainly going to continue to use things like this. Those are things that they can use for a lifetime. So those are important, those are very good to use." Teacher 1	"Just so they can make a connection with looking at a new book to um you know everything else they do in everyday life." Teacher 2
"There are kids that are three or four grades below their reading level, and then in reality we have kids that are six grades below reading level; but what can you do" Teacher 2	"They need to understand that you have to tell them about context clues, and tell them to go back and re-read passages." Teacher 2	"You want to show them that there is a reason why you did that" Teacher 2
"How do you, you just read it, come on man, come on, seriously [in relation to students not comprehending text]. And in reality they don't know what they read." Teacher 2	You don't say, "Oh, I'm going to use context clues to see what that word is." You just use it, it's written like that so that's obviously what it means… I think and that's the big thing, is just making people, making teacher aware that that's what should be happening in a good reader…Because otherwise you just sit there staring at the kid and you're like, "Why, how can you not get this?" "How is this possible?" "How	"It worked [partner reading]. Um, it wasn't…they did it, they were relatively engaged with it. They were mostly on task and it wasn't like an absolute buzz kill for them which sometimes, it's history, it's going to be boring" Teacher 2

	can you not get this?” “You speak English fine, how can you not understand this?” Teacher 2	
“We can go through and look at it word for word. We could try to explain, but I still won’t know where they’re [struggling readers] coming from.”	“Especially with history; understanding the main idea of what they’re getting at is kinda important. And then pulling out the supporting ideas, and also if you can do that when you read, then you can do that when you write.” Teacher 2	“So having something that they kind of enjoy or feel like they have some input or say in then I am fine with that” Teacher 2
“Cause they don’t understand what they are reading half of the time so what’s the point of like, you’re treading water, you aren’t getting anywhere if people don’t understand what they are reading.”	“When we were talking about comprehending your reading and how you can do that to do when you don’t understand your reading. It [specific strategy] gave me like something concrete to say, go back, look at this, you know this is what you should be doing rather than saying, why don’t you understand what you’re reading.” Teacher 2	

Appendix D

Original STERI Instrument before Expert Review

Self-Efficacy for Using Content-Area Reading Strategy Instruction

1. To what extent can you instruct students on the different reading strategies to improve comprehension?
2. To what extent can you help students choose the proper reading strategy?
3. How much do you know about using reading strategies?[5]
4. To what extent can you incorporate reading strategies into the classroom?
5. To what extent can you model successfully using reading strategies?
6. To what extent can you incorporate the reading strategies most useful to your content area?
7. How well can you explain reading strategies to struggling readers?[5]
8. How much have you been trained in using reading strategies in your classroom?[1]
9. How well can you help all students use reading strategies to improve comprehension?
10. To what extent can reading strategies be utilized in your classroom?
11. How confident are you in your ability to teach reading strategies?
12. To what extent can you demonstrate reading strategies to your students?
13. How much can you help students work through the use of different reading strategies?
14. To what extent can you meet the needs of struggling readers by using various reading strategies to improve comprehension?[4]
15. To what extent were you trained on how to utilize reading strategies?[1]
16. How much training did you receive on using reading strategies to improve comprehension?[1]
17. To what extent do you need more training in using reading strategies within your content area?[1]
18. To what extent do you wish you received more training in how to implement reading strategies in your classroom?[1]
19. To what extent can you differentiate reading strategies for different levels of readers?
20. How much can you differentiate reading strategies to improve comprehension for different students?
21. To what extent do you feel underprepared to teach reading strategies in your classroom?[1]

22. How often do you wish you know more about utilizing reading strategies within your classroom?[1]
23. How well can you demonstrate how a good reader uses reading strategies in your content area?
24. To what extent can you encourage students to find the reading strategy that works best for them?
25. To what extent can you encourage students' autonomous use of reading strategies within your classroom?
26. How familiar are you with reading strategies?[2]
27. How much do you believe reading strategies help students read in your content area?[2]
28. To what extent do you believe reading strategies should be used in your content area?[2]
29. How confident are you in your ability to teach various reading strategies to students in your content area?
30. To what extent do reading strategies help students read in your content area?[2]
31. How much can you help students monitor their own use of reading strategies?

Self-Efficacy for Teaching Struggling Readers

32. To what extent can you help struggling readers?
33. How well can you diagnose why a student is struggling in reading?
34. To what extent can you model good reading to struggling readers?
35. To what extent are you prepared to help struggling readers?[1]
36. How much have you been trained to help struggling readers?[1]
37. How confident are you in your ability to help struggling readers?
38. To what extent can you meet the needs of struggling readers?
39. How well equipped are you to help struggling readers?[1]
40. How often do you feel as though you are unprepared to meet the needs of struggling readers?[1]
41. To what extent do you need more training in meeting the needs of struggling readers?[1]
42. How confident do you feel when teaching struggling readers?
43. To what extent are struggling readers difficult to teach?
44. To what extent do you wish you received more training in instructing struggling readers?[1]
45. How confident are you in your ability to diagnose the difficulties of struggling readers?
46. To what extent were you trained to diagnose the difficulties of struggling readers?[1]
47. To what extent can you differentiate literacy instruction for struggling readers?
48. To what extent can you create intervention plans to help struggling readers?
49. How much can you differentiate course readings to help struggling readers?

50. To what extent were you trained to utilize differentiation of texts to help struggling readers?[1]
51. To what extent is it difficult to help struggling readers in your content area?[4]
52. How often do you struggle to help struggling readers in your classroom?
53. To what extent do you feel you were underprepared to meet the needs of struggling readers?[1]
54. To what extent can you model what good readers do when reading in your content area?
55. How much can you demonstrate good reading skills to help struggling readers?
56. To what extent can you encourage struggling readers to try out different reading strategies?
57. How much can you help even the most struggling readers?
58. To what extent can you adjust reading materials in your content area for struggling below-grade readers?
59. To what extent can you help struggling readers monitor their own use of reading strategies?
60. To what extent do reading strategies help struggling readers with comprehension?
61. How important is it to help struggling readers use reading strategies within your classroom?[2]

Self-Efficacy in Motivating Students to Read

62. To what extent can you motivate students to read in your content area?
63. To what extent can you motivate even the most struggling reader to read in your content area?[4]
64. How well trained are you in motivating students to read?[1]
65. To what extent do you feel ill-prepared to utilize motivational supports in your classroom?[1]
66. To what extent do you wish you received specialized training in motivating students to read in your content area?[1]
67. How much can you use differentiation as a motivation tool for students?
68. How much can you motivate students to utilize reading strategies in your classroom?[4]
69. To what extent can you motivate struggling readers in your classroom?
70. How often do you incorporate motivational supports for reading in your classroom?[4]
71. To what extent is it difficult to motivate students to read in your content area?[4]
72. To what extent do you feel underprepared to motivate students to read in your content area?[1]
73. How often do you feel ill-equipped to motivate students to read in your content area?[1]

74. To what extent can you motivate struggling readers to use reading strategies in your classroom?[4]
75. How much is motivation to read important in your content area?[2]
76. To what extent can you motivate students who are not interested in reading in your content area?
77. How confident are you in your ability to motivate all students to read in your content area?
78. How often do you use motivational supports for reading in your classroom?
79. To what extent do you believe motivation to read is important in your content area?[2]
80. To what extent can you adjust the reading materials in your content area to motivate above-grade readers?[5]
81. How much can you motivate students to care about their reading in your classroom?
82. To what extent can you use motivational supports to increase students' beliefs about their reading abilities?
83. How much can you increase students' reading beliefs using motivational supports?
84. How important is it that students' believe they can be successful readers in your content area?[2]
85. To what extent are students' beliefs about their reading ability important to their success in your classroom?[2]
86. To what extent do you wish you knew more about motivating students to be successful readers in your content area?[1]
87. To what extent is it important to motivate students to become good readers in your content area?[2]
88. How confident are you in motivating even the most struggling reader to read in your content area?[4]
89. To what extent are motivation practices a part of your teaching practice?
90. How well prepared are you to motivate even the most struggling reader in your classroom?[1]
91. How important are motivational supports in increasing students' beliefs about their reading ability in your content area?[2]

Self-Efficacy in Reading Assessment[3]

92. To what extent can you effectively assess students' comprehension in your content area?
93. How much can you adjust literacy instruction based on students' comprehension?
94. To what extent can you gauge students' comprehension based on assessments?
95. How confident are you in your ability to assess students' reading comprehension?

96. How much training did you receive in creating informal and formal assessments of students’ reading comprehension in your content area?
97. To what extent do you feel unprepared to accurately assess students’ reading comprehension in your content area?
98. How important is it to assess students’ reading comprehension in your content area?
99. How confident are you in your ability to adjust the reading strategies students’ use based on assessments?
100. To what extent can you use assessment to help students choose what reading strategies work best for them?
101. To what extent can you model how good readers use assessments to monitor their comprehension?
102. How much can you help students use assessments to monitor their comprehension in your content area?
103. To what extent can you instruct students to use assessments to gauge their comprehension?
104. To what extent can you gauge how well students’ have comprehended what they have read in your content area?
105. How important are assessments in fostering students’ beliefs about reading in your content area?
106. How comfortable are you using informal assessments to continually monitor students’ comprehension of text in your content area?
107. How important do you believe ongoing assessments are to gauge student comprehension in your content area?
108. How much can you vary your assessment types in order to most effectively monitor students’ comprehension?
109. How often do you wish you had received more specialized training in how to use assessments of reading effectively in your classroom?
110. How often do you use formal assessments of reading comprehension in your classroom?
111. How often do you use informal assessments of reading comprehension in your classroom?
112. To what extent do assessments encourage or deter students’ beliefs about their reading ability in your classroom?
113. To what extent do you feel you can effectively assess students’ reading comprehension?
114. How much can you diagnose a student’s difficulties with reading comprehension in your content area through assessments?
115. How much of a role do informal and formal assessments play in your daily teaching?
116. To what extent do assessments influence the literacy instruction within your classroom?
117. How much do you use assessments to gauge students’ comprehension of material taught in your classroom?

118. How comfortable are you creating assessments that gauge students' reading comprehension in your content area?
119. How confident are you using teacher-created assessments to determine students' level of comprehension in your content area?
120. How much do you use authentic assessments to gauge students' level of comprehension in your content area?
121. To what extent are authentic assessments important in gauging students' level of comprehension in your content area?
Note. [1]Item was removed for assessing preparation. [2]Item was removed for assessing value. [3]Entire factor and items were removed. [4]Item was removed for redundancy. [5]Item was removed for not assessing self-efficacy.

Appendix E

Revised STERI after Doctoral Literacy Student Review

Self-Efficacy for Using Content-Area Reading Strategy Instruction

1. How much can you instruct students on different reading strategies to improve comprehension?[1]
2. How confident are you in your ability to teach different reading strategies to students in your content area?
3. How confident are you in your ability to help students choose the proper reading strategy to improve their comprehension in your content area?
4. How confident are you in your ability to incorporate reading strategies into your classroom instruction? [2] (2.25; 2.8)*
5. How confident are you in your ability to model reading strategies for students in your content area?[1]
6. How confident are you in your ability to help students use reading strategies to improve comprehension in your content area?
7. How confident are you in your ability to help students use reading strategies in your classroom?[2] (2; 2.2)*
8. How confident are you in your ability to teach reading strategies in your content area?[1]
9. To what extent can you model how to use reading strategies to students in your content area?[1]
10. How confident are you in your ability to differentiate or modify reading strategies for different levels of readers in your classroom?
11. How much can you differentiate reading strategies to improve comprehension for students of all reading levels?[1]
12. How confident are you in your ability to model how an effective reader uses reading strategies in your content area?
13. How confident are you in your ability to encourage students to find the reading strategy (or strategies) that work(s) best for them in your content area?
14. How much can you encourage students' independent use of reading strategies within your classroom?[1]
15. How confident are you in your ability to help students monitor their own use of reading strategies in your content area?
16. How well can you use assessments to monitor student's use of reading strategies to increase their comprehension in your content area?[1]

17. How confident are you in your ability to assess students' use of reading strategies in your content area?

Self-Efficacy in Motivating Students to Read/in Reading

18. How confident are you in your ability to motivate students to read in your content area?
19. How confident are you in your ability to use differentiation as a motivation tool for reading in your content area?
20. How confident are you in your ability to motivate students to improve their comprehension when reading in your content areas?
21. To what extent do you incorporate motivation for reading in your classroom?[1]
22. How confident are you in your ability to motivate difficult students to read in your content area?
23. How confident are you in your ability to motivate students who are not interested in reading in your content area?
24. How confident are you in your ability to motivate all students to read in your content area?
25. How confident are you in your ability to motivate students to care about their reading in your classroom?
26. How confident are you in your ability to use motivational supports to increase students' beliefs about their reading abilities in your content area?
27. How confident are you in your ability to use various types of assessments to motivate students in reading in your content area?

Self- Efficacy for Teaching Struggling Readers

28. How confident are you in your ability to diagnose why a student is struggling with reading in your content area?
29. How confident are you in your ability to model effective reading in your content area to help struggling readers?
30. How confident are you in your ability to help struggling readers comprehend text in your content area?
31. How confident do you feel when teaching struggling readers in your content area?[1]
32. How confident are you in your ability to meet the reading needs of struggling readers in your content area?[2] (2.23; 3)*
33. How confident are you in your ability to help struggling readers in your content area?
34. How confident are you in your ability to diagnose the difficulties of struggling readers?
35. How confident are you in your ability to differentiate (or change) literacy instruction for struggling readers in your content area?

36. How confident are you in your ability to create intervention plans to help struggling readers in your content area?

37. How confident are you in your ability to encourage struggling readers to improve their comprehension when reading in your content area?[2] (2.25; 2.6)*

38. How confident are you in your ability to differentiate course readings to help struggling readers in your content area?

39. How confident are you in your ability to model effective reading skills to help struggling readers in your content area?

40. How much can you encourage even the most struggling reader to read in your content area?[2] (2.25; 6)*

41. How confident are you in your ability to adjust reading materials in your content area for struggling readers?

42. How confident are you in your ability to help struggling readers assess their comprehension in your content area?

43. How confident are you in your ability to explain model the importance of comprehension to struggling readers in your content area?

44. How confident are you in your ability to help struggling readers comprehend texts in your content area?[1]

45. How confident are you in your ability to assess struggling readers' difficulties in order to improve their reading in your content area?

46. How confident re you in your ability to use a variety of assessments to help struggling readers in your content area?

Note. [1]*Item was removed for redundancy.* [2]*Item was removed for low clarity/relevance. *Mean ratings for (clarity; relevance).*

Appendix F

Semi-Structured Cognitive Interview Questions

Anticipated Probes, asked of all participants

Initial Probes (asked for every question):

1. Can you please read the item aloud?

2. Can you put the question into your own words?

3. If you responded to this question what score would you choose?

4. Why would you respond in that way?

Further Probes (asked once per participant):

1. What do you think of when you hear the term reading strategy?

2. What do you think of when you hear the term struggling readers?

3. What do you think of when you hear the term motivation to read?

Conditional Probes, asked of some participants

1. It sounded like it was difficult for you to paraphrase that question; what made the item difficult to paraphrase?

2. You seemed to really struggle when answering that question; why do you think that is?

3. Why do you think you had difficulty defining that term (e.g., struggling readers)?

Appendix G

Informed Consent form for Cognitive Interviews

RESEARCH PROCEDURES
This research is being conducted to gain insights in teachers' beliefs about reading instruction across the content areas. If you agree to participate in this study, you will be asked to complete a survey while answering interview questions about the meaning and clarity of items on the survey. The interview will last approximately 30-45 minutes. With your permission, the interview will be audio-recorded in order to accurately capture your responses; however, your name will never be used and no one but the primary researcher will hear the recording.

RISKS
There are no foreseeable risks for participating in this research.

BENEFITS
There are no benefits to you as a participant other than providing research that will help advance the teacher preparation and reading instruction.

CONFIDENTIALITY
The data in this study will be confidential. No identifying information will be collected at any time during the study. With your permission, the interview will be audio-recorded. All information obtained from the interview will be used for general revisions to the measure, not to report on individuals' responses.

PARTICIPATION
Your participation is voluntary, and you may withdraw from the study at any time and for any reason. If you decide not to participate or if you withdraw from the study, there is no penalty or loss of benefits to which you are otherwise entitled. There are no costs to you or any other party.

CONTACT
This research is being conducted by Erin Ramirez, a PhD student, at George Mason University. If you have any questions or research-related issues Erin may be reached at xxx-xxx-xxxx. You may also contact Dr. Michelle Buehl, Associate Professor of Education at George Mason University and the principal investigator, at xxx-xxx-xxxx. You may contact the George Mason University Office of Research Subject Protections at

703-993-4121 if you have questions or comments regarding your rights as a participant in the research. This research has been reviewed according to George Mason University procedures governing your participation in this research.

Appendix H

Demographic Teacher Survey

What subject area do you teach?
Social Studies
English
Math
Science

Highest level of education? Check all that apply
Bachelors
Bachelors + additional coursework
Masters
Masters + additional coursework
Reading Specialist Degree
Doctorate

Number of literacy courses taken:
Number of literacy professional development sessions attended:
Years of high school teaching experience:

Gender:
Male
Female

Ethnicity/Race:
African-American/Black
Hispanic/Latino
Caucasian
Asian
American Indian/Alaska Native
Hawaiian Native/Pacific Islander
Multiracial
Other:

State where you teach: _______________
Demographics of school district:

Appendix I

STERI Form A for Initial Cognitive Interviews

Directions: Thinking about your primary teaching assignment, please indicate your responses to each of the following statements. All answers will be kept confidential.	**In your content area, how well can you…**					
1. Model effective reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
2. Modify instruction to motivate students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
3. Help students monitor their use of reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
4. Encourage students to find the reading strategy (or strategies) that work(s) best for them?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
5. Change your instruction to meet the needs of individual readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
6. Model effective reading for struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

7. Assess students' use of reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
8. Motivate students to improve their understanding when reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
9. Determine the reading needs of struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
10. Create intervention plans to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
11. Motivate students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
12. Determine why a student is having difficulty with reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
13. Encourage students to self-assess their reading comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
14. Encourage uninterested students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
15. Inspire students of all reading abilities to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
16. Help struggling readers comprehend text?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

Appendix J

STERI Form B for Initial Cognitive Interviews

Directions: Thinking about your primary teaching assignment, please indicate your responses to each of the following statements. All answers will be kept confidential.	**In your content area, how well can you…**					
1. Teach different reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
2. Emphasize the importance of reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
3. Help students apply reading strategies to improve comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
4. Use various types of assessments to motivate students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
5. Use a variety of assessments to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
6. Help students choose which reading strategy to use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
7. Increase students' confidence in reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

8. Select course readings to match the needs of individual readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
9. Communicate the importance of comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
10. Adapt reading materials for struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
11. Identify improper reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
12. Differentiate reading strategy instruction?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
13. Help struggling readers self-assess their own comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
14. Utilize a variety of texts to encourage students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
15. Help struggling readers improve their comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

Appendix K

STERI Form C for Second Round of Cognitive Interviews

Directions: Thinking about your primary teaching assignment, please indicate your responses to each of the following statements. All answers will be kept confidential.	**In your content area, how well can you…**					
2. Teach different reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
3. Help struggling readers comprehend text?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
4. Identify improper reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
5. Change your instruction to meet the needs of readers of different abilities?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
6. Create intervention plans to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
7. Select course readings to match the needs of individual students?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
8. Communicate the importance of readings comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

9.	Help struggling readers improve their comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
10.	Modify instruction to encourage students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
11.	Encourage uninterested students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
12.	Motivate students to read about course topics?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
13.	Use various types of assessments to motivate students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

Appendix L

Informed Consent Form for Survey Participation

RESEARCH PROCEDURES

This research is being conducted to gain insights about teachers' beliefs about reading instruction across the content areas. If you agree to participate in this study, you will be asked to complete one survey on teacher beliefs about reading instruction. You will complete this survey online and it will take approximately 20-30 minutes to complete.

RISKS

There are no foreseeable risks for participating in this research.

BENEFITS

There are no benefits to you as a participant other than providing research that will help advance the field of teacher preparation and reading instruction.

CONFIDENTIALITY

The data in this study will be confidential. No identifying information will be collected at any time during the study. The entire study will be conducted online and you will not be asked to provide any follow-up information at any time.

PARTICIPATION

Your participation is voluntary, and you may withdraw from the study at any time and for any reason. If you decide not to participate or if you withdraw from the study, there is no penalty or loss of benefits to which you are otherwise entitled. There are no costs to you or any other party.

CONTACT

This research is being conducted by Erin Ramirez, a PhD student, at George Mason University. If you have any questions or research-related issues Erin may be reached at xxx-xxx-xxxx. You may also contact Dr. Michelle Buehl, Associate Professor of Education at George Mason University and the principal investigator, at xxx-xxx-xxxx. You may contact the George Mason University Office of Research Subject Protections at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research. This research has been reviewed according to George Mason University procedures governing your participation in this research.

☐ I have read and understand the aforementioned and willingly provide my consent.

Appendix M

Convergent Validity Subscales from the Teacher Sense of Efficacy Scales
(TSES; Tschannen-Moran and Woolfolk Hoy, 2001)

Efficacy for Student Engagement (α = .81)

2. How much can you do to motivate students who show low interest in school work?
3. How much can you do to get students to believe they can do well in school work?
4. How much can you do to help your students value learning?
11. How much can you assist families in helping their children do well in school?

Efficacy for Instructional Strategies (α = .86)

5. To what extent can you craft good questions for your students?
9. How much can you use a variety of assessment strategies?
10. To what extent can you provide an alternative explanation or example when students are confused?
12. How well can you implement alternative strategies in your classroom?

Note. Item numbers are from original scale.

Appendix N

Convergent Validity Subscales from Goal Orientation for Teaching
(Butler 2012, Butler & Shibaz, 2014).

Mastery Goals (α = .74)

1. I'd feel I had a successful day in school if I learned something new about myself as a teacher.
2. I'd feel I had a successful day in school if something that happened in class made me want to deepen my professional knowledge.
3. I'd feel I had a successful day in school if a student asked a question in class that made me think again about the subject matter.
4. I'd feel I had a particularly successful day in school if I saw that I am developing professionally and teaching more effectively than in the past.

Relational Goals (α = .86)

1. My main goal as a teacher is to build a deep personal relationship with each and every student
2. More than anything, I strive to create and maintain meaningful relationships with students.
3. My relationships with students are more important to me than anything else in my role as teacher.
4. I feel that I have succeeded as a teacher if I create close and warm relationships with students and classes.

Appendix O

Discriminant Validity Subscales from Invitations of Self and Others Scale
(Morris, 2010)

Invitations of Self ($\alpha = .89$)

IS1: I am quick to recognize my own value as a teacher.
IS2: I plan time for enjoyable activities that I can do on my own.
IS3: I congratulate myself on my teaching successes.
IS4: I pay attention to my own needs.
IS5: I forgive myself for my mistakes in teaching.
IS6: I am impressed with my own teaching abilities.

Invitations of Others ($\alpha = .73$)

IO1: I like to include others in the activities in my classroom.
IO2: I try not to be critical of the people with whom I work.
IO3: I congratulate my colleagues on their successes.
IO4: I forgive my colleagues for their mistakes.
IO5: I am impressed with the abilities of others with whom I work.
IO6: I am quick to recognize the value of others with whom I work.
Note. Numbers are from original scale.

Appendix P

Recruitment Letter to Supervisors and Teachers

To Whom it May Concern,

My name is Erin Ramirez and I am a doctoral candidate at George Mason University. As a former secondary teacher, I am very interested in how teacher education and professional development can affect teacher beliefs and practices.
I am writing to request your content area teachers' participation in my dissertation study looking at teacher beliefs about literacy instruction. The goal of the dissertation is to capture teacher beliefs about literacy instruction in the content areas to foster personalized professional development opportunities for secondary teachers.
If you could please forward the following email and survey link to your teachers it would be greatly appreciated. The survey will take 20-30 minutes to complete and will be completely anonymous, with no identifying information provided.
I appreciate your help in this process, and am available at eramire4@gmu.edu if you have any questions.

Teachers,
I need your help! As a former secondary teacher, I always struggled to implement literacy instruction into my classroom. I wished I was better prepared and had more targeted professional development to combat my struggles. Now I'm a doctoral candidate at George Mason University and want to use my dissertation to capture how practicing teachers feel about literacy instruction in order to foster more personalized teacher professional development opportunities for literacy instruction.
If you could PLEASE take 20-30 minutes out of your busy day to take the following survey: (provide link when available) it would be GREATLY APPRECIATED! From experience, I know how busy teaching is but from one former teacher to another, you taking this survey would be tremendously helpful for my PhD studies and more importantly for the field to help design professional development opportunities more catered to teacher's needs. No identifying information will be obtained and your responses will be completely anonymous. The more teachers participate in the survey, the greater the impact of the research; so if you can forward this message on to any other secondary content area teachers you know it would be greatly appreciated.
Thank you so very much for your time and professional courtesy. If you have any questions, feel free to contact me at eramire4@gmu.edu.

Appendix Q

Quantitative Data from Second Round of Expert Review

	Items	Clarity	Relevance
1.	How confident are you in your ability to teach different reading strategies?	3.33	3.67
2.	How confident are you in your ability to help students choose the proper reading strategy to improve their comprehension in your content area?	2.67*	3.67
3.	How confident are you in your ability to help students use reading strategies to improve comprehension in your content area?	3.50	4.00
4.	How confident are you in your ability to differentiate or modify reading strategies for different levels of readers in your classroom?	3.00	3.67
5.	How confident are you in your ability to model how an effective reader uses reading strategies in your content area?	3.00	3.67
6.	How confident are you in your ability to encourage students to find the reading strategy (or strategies) that work(s) best for them in your content area?	3.33	3.33
7.	How confident are you in your ability to help students monitor their own use of reading strategies in your content area?	3.33	3.67
8.	How confident are you in your ability to assess students' use of reading strategies in your content area?	3.33	4.00
9.	How confident are you in your ability to diagnose why a student is struggling with reading in your content area?	3.33	4.00
10.	How confident are you in your ability to model effective reading in your content area to help struggling readers?	3.33	4.00
11.	How confident are you in your ability to help struggling readers comprehend text in your content area?	3.67	4.00
12.	How confident are you in your ability to help struggling readers in your content area?	2.33*	2.33*
13.	How confident are you in your ability to diagnose the difficulties of struggling readers?	3.67	3.33
14.	How confident are you in your ability to differentiate (or change) literacy instruction for struggling readers in your content area?	3.33	4.00
15.	How confident are you in your ability to create intervention plans to help struggling readers in your content area?	3.33	4.00

16. How confident are you in your ability to differentiate course readings to help struggling readers in your content area?	3.33	4.00
17. How confident are you in your ability to model effective reading skills to help struggling readers in your content area?	3.00	3.67
18. How confident are you in your ability to adjust reading materials in your content area for struggling readers?	3.00	4.00
19. How confident are you in your ability to help struggling readers assess their comprehension in your content area?	3.00	4.00
20. How confident are you in your ability to model the importance of comprehension to struggling readers in your content area?	2.67*	3.00
21. How confident are you in your ability to assess struggling readers' difficulties in order to improve their reading in your content area?	2.33*	4.00
22. How confident re you in your ability to use a variety of assessments to help struggling readers in your content area?	3.33	4.00
23. How confident are you in your ability to motivate students to read in your content area?	3.33	4.00
24. How confident are you in your ability to use differentiation as a motivation tool for reading in your content area?	3.33	3.67
25. How confident are you in your ability to motivate students to improve their comprehension when reading in your content areas?	3.67	3.67
26. How confident are you in your ability to motivate difficult students to read in your content area?	2.67*	4.00
27. How confident are you in your ability to motivate students who are not interested in reading in your content area?	3.00	3.67
28. How confident are you in your ability to motivate all students to read in your content area?	3.00	3.67
29. How confident are you in your ability to motivate students to care about their reading in your classroom?	2.67*	3.33
30. How confident are you in your ability to use motivational supports to increase students' beliefs about their reading abilities in your content area?	2.33*	3.67
31. How confident are you in your ability to use various types of assessments to motivate students in reading in your content area?	3.33	4.00

Note. *= Mean lower than 3.00

Appendix R

Qualitative Data from Second Round of Expert Review

Directions: Thinking about your primary teaching assignment, please indicate your responses to each of the following statements. All answers will be kept confidential.				
Item	**Final Items**	**Reviewer Qualitative Notes**		
Stem changed from “How confident are you” to “In your content area, how well can you”				
1. How confident are you in your ability to teach different reading strategies to students in your content area?	Teach different reading strategies?			Omit “to students”
2. How confident are you in your ability to help students choose the proper reading strategy to improve their comprehension in your content area?	Help students choose which reading strategy to use?		Might want to broaden this a bit; “the” proper reading strategy implies there is a correct “one” to choose. Perhaps consider the purpose for which the student would select the strategy?	How well can you help students choose the proper reading comprehension strategy? Or, How certain are you that you can …

3. How confident are you in your ability to help students use reading strategies to improve comprehension in your content area?	Help students apply reading strategies to improve comprehension?			This seems redundant with the previous.
4. How confident are you in your ability to differentiate or modify reading strategies for different levels of readers in your classroom?	Differentiate reading strategy instruction?			Double barreled item
5. How confident are you in your ability to model how an effective reader uses reading strategies in your content area?	Model effective reading strategy use?			How well can you model effective reading strategy use?
6. How confident are you in your ability to encourage students to find the reading strategy (or strategies) that work(s) best for them in your content area?	Encourage students to find the reading strategy (or strategies) that work(s) best for them?	This is a bit close to #2	This one overlaps with #2, but is clearer in that it references the selection of multiple strategies.	
7. How confident are you in your ability to help students monitor their own use of reading strategies in your content area?	Help students monitor their use of reading strategies?			
8. How confident are you in your ability to assess students' use of reading strategies in your content area?	Assess students' use of reading strategies?			

9. How confident are you in your ability to diagnose why a student is struggling with reading in your content area?	Determine why a student is having difficulty with reading?		The “diagnose” sounds like something within the realm of reading specialists.	See above. Alternative: How certain are you that you can diagnose …
10. How confident are you in your ability to model effective reading in your content area to help struggling readers?	Model effective reading for struggling readers?			How is this different from #17 below?
11. How confident are you in your ability to help struggling readers comprehend text in your content area?	Help struggling readers comprehend text?			
12. How confident are you in your ability to help struggling readers in your content area?	Item Deleted	Seems very general—I would let the more specific items carry the day	Too vague; help to do what?	This might not have anything to do with literacy instruction. “To help” is pretty vague.
13. How confident are you in your ability to diagnose the difficulties of struggling readers?	Determine the reading needs of struggling readers?	How is this different from #9?	Overlaps with #9	
14. How confident are you in your ability to differentiate (or change) literacy instruction for struggling readers in your content area?	Change your instruction to meet the needs of individual readers?			Choose differentiate or change

15. How confident are you in your ability to create intervention plans to help struggling readers in your content area?	Create intervention plans to help struggling readers?			
16. How confident are you in your ability to differentiate course readings to help struggling readers in your content area?	Select course readings to match the needs of individual readers?			
17. How confident are you in your ability to model effective reading skills to help struggling readers in your content area?	Item Deleted	How is this different from #10?	Overlaps with #10	
18. How confident are you in your ability to adjust reading materials in your content area for struggling readers?	Adapt reading materials for struggling readers?	Are you getting at something different from #s 14 and 16?	Overlaps with #14	
19. How confident are you in your ability to help struggling readers assess their comprehension in your content area?	Help struggling readers self-assess their own comprehension?		Are you referring to a student's ability to self-assess?	
20. How confident are you in your ability to model the importance of comprehension to	Communicate the importance of comprehension?		I'm not sure what criteria teachers would use to know whether they had "modeled the	I don't think teachers model importance.

struggling readers in your content area?			importance of comprehension."	
21. How confident are you in your ability to assess struggling readers' difficulties in order to improve their reading in your content area?	Help struggling readers improve their comprehension?	How is this different from #9 and #13?		Double barreled.
22. How confident re you in your ability to use a variety of assessments to help struggling readers in your content area?	Use a variety of assessments to help struggling readers?		Might consider specifying formal v. informal assessment, or formative v. summative assessments	
23. How confident are you in your ability to motivate students to read in your content area?	Encourage students to self-assess their reading comprehension?			
24. How confident are you in your ability to use differentiation as a motivation tool for reading in your content area?	Modify instruction to motivate students to read?			Do teachers know what this means? (I'm not sure I do.)
25. How confident are you in your ability to motivate students to improve their comprehension when reading in your content areas?	Motivate students to improve their understanding when reading?			I don't think this is particularly clear. To motivate to improve one's understanding? Maybe …
26. How confident are you in your ability to motivate	Inspire students of all reading abilities to read?	Do you mean students who have	Define "difficult students"	You might specify what you mean by "difficult students"

difficult students to read in your content area?		behavior problems or who have language problems, or who lack vocabulary……?		
27. How confident are you in your ability to motivate students who are not interested in reading in your content area?	Encourage uninterested students to read?	This is clearer.		Hmmmm. Motivate them to do what?
28. How confident are you in your ability to motivate all students to read in your content area?	Motivate students to read?	This may be a bit general	"Motivate all students" is pretty broad. Are you referring to engagement or to applying content area reading outside of the classroom?	
29. How confident are you in your ability to motivate students to care about their reading in your classroom?	Emphasize the importance of reading?		Not sure what this question is asking. See comment for #28.	To motivate them to care? I'm wondering how teachers will understand this. I think it's worth giving some of these items to teachers and doing a think aloud.
30. How confident are you in your ability to use	Increase students' confidence in reading?		What is meant by "increase students'	What does it mean to increase beliefs?

motivational supports to increase students' beliefs about their reading abilities in your content area?			beliefs?" Build their confidence?	What kind of beliefs? Self-doubt is a belief about one's ability, but I don't think that's what you're after here.
31. How confident are you in your ability to use various types of assessments to motivate students in reading in your content area?	Use various types of assessments to motivate students to read?		As with question #22, you may wish to specify certain types of assessments.	
32. Added Item	Utilize a variety of texts to encourage students to read?			
33. Added Item	Identify improper reading strategy use?			

Appendix S

Revised STERI after Second Round of Expert Review

Directions: Thinking about your primary teaching assignment, please indicate your responses to each of the following statements. All answers will be kept confidential.	**In your content area, how well can you…**					
1. Teach different reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
2. Assist students' in choosing which reading strategy to use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
3. Help students apply reading strategies to improve comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
4. Incorporate reading strategy instruction?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
5. Model effective reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
6. Encourage students to find the reading strategy (or strategies) that work(s) best for them?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
7. Help students monitor their use of reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

8. Assess students' use of reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
9. Determine why a student is having difficulty with reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
10. Model effective reading for struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
11. Help struggling readers comprehend text?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
12. Determine the reading needs of struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
13. Change your instruction to meet the needs of individual readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
14. Create intervention plans to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
15. Select course readings to match the needs of individual readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
16. Adapt reading materials for struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
17. Help struggling readers self-assess their own comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
18. Communicate the importance of comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
19. Help struggling readers improve their comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
20. Use a variety of assessments to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

21. Encourage students to self-assess their reading comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
22. Modify instruction to motivate students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
23. Motivate students to improve their understanding when reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
24. Influence students of all reading abilities to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
25. Encourage uninterested students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
26. Motivate students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
27. Cause students to value reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
28. Increase students' confidence in reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
29. Use various types of assessments to motivate students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
30. Utilize a variety of texts to encourage students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
31. Identify improper reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

Appendix T

Revised STERI after Initial Round of Cognitive Interviews

Items	**Content Area of Teacher**								**Revised Items**
"In your content area, how well can you…"	Social Studies		Mathematics		Science		English		
Form A	YRS: 3	G: Male	YRS: 1	G: Male	YRS: 11	G: Male	YRS: 22	G: Female	
1. Model effective reading strategy use?	"This question is very clear…can I model correct reading strategy use"		"I don't do it, but yes I understand what you're asking me here."		"I understand the question, and I would say very little because I often just refer them to the toolbox, I'm not modeling myself very often."		"Yes, you're asking if I think I can effectively model how to use reading strategies."		*No change*
2. Modify instruction to motivate students to read?	"Yes, I modify my instruction as much as I can, but it is very time consuming."		"Since I don't focus on reading in my classroom I don't change my instruction to deal with		"Yes I'm always differentiating my instruction." Follow up: "So you're changing your instruction to		"You have to constantly change up your instruction to keep them [students] engaged" Follow up: "Are you modifying		*Modify instruction to encourage students to read*

	Follow up: "Yes, but do you modify instruction to motivate students to read?" Teacher: "Well if I'm switching up instruction I should be motivating them to learn."	reading. In math I do, but not in reading."	encourage them to read?" Teacher: "Well I differentiate with my reading materials and how I teach to help get the grade."	your instruction to motivate students to read?" Teacher: "Yes, like I said if you don't change up instruction than students are not engaged."	
3.					*No change*
4.					*No change*
6. Change your instruction to meet the needs of individual readers?	Reworded question: "Do I change my instruction to help all my students"	"Yes, what works for some will definitely not work for others; especially in my special education classes.""	"Because it's content specific to science so it is extremely important that they read and understand what we're reading, so I have to mix things up."	"They [students] all need something a little bit different and so I'm constantly changing up what I'm doing and what I'm giving them to read to help them get the content."	*Change your instruction to meet the needs of readers of different abilities?*
7. Assess students' use of reading strategies?	"I know what the question is asking, do I know when they are using a strategy incorrectly, but I'm not even sure how to go about doing that, that's not my				*No change*

	forte, I don’t feel strong in that area.”				
8. Motivate students to improve their understanding when reading?	“I’m not sure any teacher can really do this. I can tell them over and over again to read the textbook but that doesn’t mean they’ll do it. I’m not sure I really have any influence in this area.”			“If I could do that I would be a better teacher” follow up “so do you think the question is difficult because of the wording, or it’s asking too many things, or is it something else” teacher “For me, maybe being a long-time teacher, just the answer itself is difficult…motivating students is just a bear of a topic”	*No change*
9. Determine the reading needs of struggling readers?	“To me, that’s a difficult one because I’m in a social studies classroom and so I’m working on the content and this seems to me like something the reading specialist should be doing.”	“That’s what a specialist is for” Follow up: “do you use the reading specialist?” “I don’t know if we have one anymore.”	“I don’t know if that really applies to me. I mean they have the strategy toolbox in the front of the book to use if they need help.”	Reworded the question: “Can I figure out what a struggling reader needs to be successful.”	*No change*

10. Create intervention plans to help struggling readers?	"This goes back to the reading specialist. I'm not equipped to make any type of plans for students; that'd be 100+ plans. I don't have time for that."	"I understand the question, but in no way do I see this as part of my job. Students should know how to read but that's what English is for."	"I am not sure if I'd even know how to start this process. There's a toolbox in the front of the book that I have them go to when they're confused."	"This one is tough because our reading specialist is in charge of creating intervention plans. But, I do know how to work with her and how to use them."	*No change*
11. Motivate students to read?	"Are you talking about academic reading, what type of reading? For enjoyment? For comprehension?"	"I'm not sure what you're really asking me here."	"This seems vague. Motivate them to read to get the grade?"	"Motivate them to become lifelong learners? Intrinsically motivate them? Can any teacher really motivate a student to do anything?"	*Motivate students to read about course topics?*
12. Determine why a student is having difficult with reading?	"Again, I'm not sure how I would answer this one, it's not something I really think about. We have the book and we have to use it."	"I don't really teach reading in my class; it's math."	"I don't concentrate so much on reading because I know that they do that in English; figure they'll get it somewhere."	Reworded the question: "How well can I really troubleshoot why a student is having a hard time reading."	*No change*
13.					*No change*
14. Encourage uninterested students to read?	"I think this is simple and to the point; I don't know how to reword it other than to say, can I get	Reworded question: "Can I get students who don't care to read."	"I understand what you're getting at; here I think I'd answer quite a bit because I'm always trying	"This is very clear; can I get students who don't care about English to read. If I could do this my life	*No change*

	uninterested students to read."		to find them other science related things to read to try and get them into the content."	would be a lot easier (laughs)."	
15. Inspire students of all reading abilities to read?	"Yes, I know what you're asking with this question…can I get all students to read."	"Hmmm…I think it's somewhat to very little [referring to anchor scores], in math I don't really think about their reading ability."	"I'd say quite a bit because they have to read to get the grade and so I'm always giving them different types of things about science to read to get them to want to read for understanding."	"I do this a lot because in English we have no choice, they have to read and so I'm constantly pushing them to want to read no matter the level, they have to read in order to succeed."	*No change*
16.					
Form B	**Social Studies**	**Mathematics**	**Science**	**English**	**Revised Items**
"In your content area, how well can you…"	YRS: 9 / G: Female	YRS: 20 / G: Female	YRS: 3 / G: Female	YRS: 4 / G: Female	
1. Teach different reading strategies?	"This is an easily understood question, and even though I don't teach reading strategies daily, I think I could moderately teach them effectively."	Rewording of question: "Do I teach different reading strategies in my classroom"	"This is a question I understand, but I don't teach reading strategies in my classroom. I offer different types of reading materials, but not strategies explicitly."	"This is very clear and something I do often. This is English, if they can't read it than they're doomed. I don't teach strategies per say but I teach them how to be successful reading	*No change*

				and what I do when I get stuck."	
2. Emphasize the importance of reading?	"I can't even rephrase this without just saying the same thing again, but yes I emphasize reading quite a bit in my classroom."	Rewording of question: "Can I stress to students how important it is to read"	"Yes, I tell them all the time how important it is to read in science because labs, tests, and everything is text related. So yes I'm always telling them to read."	"This is super clear. I do this all the time. Again it's English, they have to read because everything comes from the text."	*No change*
3.					*No change*
4.					*No change*
5. Use a variety of assessments to help struggling readers?	"I would say moderately to quite a bit because I'm always trying to change up my assessments to give students of all abilities to opportunity to do well."	Rewording of question: "You're asking if I change of my instruction for all types of students, even those who struggle reading."	"Yes, I am often changing the types of assessments I'm giving to help all students succeed in science. It's not always a reading issue, more an understanding of the content which could be related to reading or not."	"Yes, and especially when we begin a new topic I'll do things that are related and are more creative to motivate them; so I'll go very personal with some assessments and then work towards a more material driven approach."	*No change*
6. Help students choose which reading	"This question is very clear, do I help students choose the reading	Rewording of question: "Do I help students choose	"I don't particularly do that with each individual student	"Yes, and I also work with them on choosing different strategies at different	*No change*

strategy to use?	strategy to use (laughs); sorry I didn't mean to rephrase it exactly like it's written but it's clear."	reading strategies?"	because I'm not a reading classroom, I may mention a way to help them understand but I'm not helping them choose strategies per say."	times, like sometimes you need to use prediction but other times that won't work. So I try to help them see there's more than one strategy they can use."	
7.					*No change*
8. Select course readings to match the needs of individual readers?					*Select course readings to match the needs of individual students?*
9. Communicate the importance of comprehension?	"I find this question confusing, Comprehension for what?"	Rewording of item: "Do I tell students how important comprehension is. I don't really deal in terms of reading in my class, but I tell them math is important."	"Is this just them knowing it [the material]? The question is somewhat vague."	"Comprehension of texts, of reading strategies, of what? I'm not sure what you're asking me here?	*Communicate the importance of reading comprehension?*
10.					*No change*
11. Identify improper reading strategy use?	"I don't know that I could necessarily know right away	Rewording of question:	"I don't focus on reading strategies in my classroom	"This is clear and I understand the question, but since I	*No change*

	that a student was using a strategy wrong, but I would probably see it in their writing and inability to do well on tests”	“Do I know if a student is using a reading strategy wrong? Honestly I know I cannot do this, maybe very little if that.”	because I teach science, so I think I do this very little to not at all.”	don’t have formal training in reading strategies I don’t know that I’d know if a student was using one wrong.”	
12.					*No change*
13.					*No change*
14.					*No change*
15.					*No change*

Appendix U

Finalized STERI after Second Round of Cognitive Interviews

Directions: This survey is designed to examine how you view reading in your content area. There are no right or wrong answers to any of the questions. Please consider your views of reading in your content area and answer each question honestly. All answers will be kept confidential.	**In your content area, how well can you…**					
1. Model effective reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
2. Modify instruction to encourage students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
3. Help students monitor their use of reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
4. Encourage students to find the reading strategy (or strategies) that work(s) best for them?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
5. Change your instruction to meet the needs of readers of different abilities?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
6. Model effective reading for struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

7. Assess students' use of reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
8. Motivate students to improve their comprehension when reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
9. Determine the reading needs of struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
10. Create intervention plans to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
11. Motivate students to read about course topics?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
12. Determine why a student is having difficulty with reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
13. Encourage students to self-assess their reading comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
14. Inspire students of all reading abilities to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
15. Help struggling readers comprehend text?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
16. Teach different reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
17. Emphasize the importance of reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
18. Help students apply reading strategies to improve comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
19. Use various types of assessments to	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

	motivate students to read?						
20.	Encourage uninterested students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
21.	Use a variety of assessments to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
22.	Help students choose which reading strategy to use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
23.	Increase students' confidence in reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
24.	Use a variety of texts to encourage students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
25.	Select course readings to match the needs of individual students?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
26.	Communicate the importance of reading comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
27.	Adapt reading materials for struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
28.	Identify improper reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
29.	Differentiate reading strategy instruction?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
30.	Help struggling readers self-assess their comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
31.	Help struggling readers improve their comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

Appendix V

Finalized Survey Package

INFORMED CONSENT FORM FOR SURVEY PARTICIPATION

RESEARCH PROCEDURES

This research is being conducted to gain insights about teachers' beliefs about reading instruction across the content areas. If you agree to participate in this study, you will be asked to complete one survey on teacher beliefs about reading instruction. You will complete this survey online and it will take approximately 15-20 minutes to complete.

RISKS

There are no foreseeable risks for participating in this research.

BENEFITS

There are no benefits to you as a participant other than providing research that will help advance the field of teacher preparation and reading instruction.

CONFIDENTIALITY

The data in this study will be confidential. No identifying information will be collected at any time during the study. The entire study will be conducted online and you will not be asked to provide any follow-up information at any time. While it is understood that no computer transmission can be perfectly secure, reasonable efforts will be made to protect the confidentiality of your transmission.

PARTICIPATION

Your participation is voluntary, and you may withdraw from the study at any time and for any reason. If you decide not to participate or if you withdraw from the study, there is no penalty or loss of benefits to which you are otherwise entitled. There are no costs to you or any other party.

CONTACT

This research is being conducted by Erin Ramirez, a PhD student, at George Mason University. If you have any questions or research-related issues Erin may be reached at xxx-xxx-xxxx. You may also contact the Co-Principal Investigators: Dr. Michelle Buehl, Associate Professor of Education at George Mason University at xxx-xxx-xxxx or Dr. Angela Miller, Assistant Professor of Education at George Mason University at xxx-xxx-xxxx. You may contact the George Mason University Office of Research Integrity & Assurance at (703) 993-4121 if you have questions or comments regarding your rights as a participant in the research. This research has been reviewed according to George Mason University procedures governing your participation in this research.

Office of Research Integrity & Assurance
IRB: For Official Use Only
Project Number: 682512-1
Page 1 of 1

Do you agree to the above terms?
By clicking Yes, you consent that you are willing to answer the questions in this survey.
Yes No

1. What is your primary teaching assignment (the majority of classes you teach)?

Social Studies English Math Science
Other:________ (If a participant chooses other, they will be redirected to a page that tells them they are do not meet the requirements to participate in the study)

2. Years teaching in this teaching assignment: ______________

Thinking about the students in your classes within this teaching assignment:

3. How many classes do you teach in this assignment?

4. What grade levels do you teach in your primary teaching assignment (check all that apply)?

9th 10th 11th 12th

5. Roughly, what is the percent of students who are English Language Learners (ELLs):

0-25% 26-50% 51-75% 75% or more

6. Roughly, what is the percent of students who have Individualized Learning Plans (IEPs):

0-25%　　26-50%　　51-75%　　75% or more

7. Roughly, what percent of students are above-grade learners?

0-25%　　26-50%　　51-75%　　75% or more

8. Roughly, what percent of students are on-grade learners?

0-25%　　26-50%　　51-75%　　75% or more

9. Roughly, what percent of students are below-grade learners?

0-25%　　26-50%　　51-75%　　75% or more

10. If applicable, what is your secondary teaching assignment:

Social Studies　　English　　Math　　Science
Other: ________________

11 Your highest level of education:

Bachelors　　Bachelors + additional coursework　　Masters
Masters + additional coursework　　Reading Specialist　　Math Specialist　　Education Specialist Degree　　Doctorate (Ed.D/Ph.D)
Other:_________________

12. Number of literacy courses taken for credit:______________

13. Number of literacy professional development sessions attended (e.g., in-service training, summer workshops, professional or school conferences): ________________

14. Gender:

Male　　Female

15. Ethnicity/Race (select all that apply):

African-American/Black Hispanic/Latino Caucasian Asian
American Indian/Alaska Native Hawaiian Native/Pacific Islander
Multiracial Other:

16. State where you teach: _______________

17. Do you teach in a Title I school?

Yes No

18. Total years of teaching experience: _______________

Directions: This survey is designed to examine how you view reading in your content area. There are no right or wrong answers to any of the questions. Please consider your views of reading in your content area and answer each question honestly. All answers will be kept confidential.

	In your content area, how well can you…					
19. Model effective reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
20. Modify instruction to encourage students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
21. Help students monitor their use of reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
22. Encourage students to find the reading strategy (or strategies) that work(s) best for them?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
23. Change your instruction to meet the needs of readers of different abilities?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
24. Model effective reading for struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

25.	Assess students' use of reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
26.	Motivate students to improve their comprehension when reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
27.	Determine the reading needs of struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
28.	Create intervention plans to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
29.	Motivate students to read about course topics?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
30.	Determine why a student is having difficulty with reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
31.	Encourage students to self-assess their reading comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
32.	Inspire students of all reading abilities to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
33.	Help struggling readers comprehend text?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
34.	Teach different reading strategies?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
35.	Emphasize the importance of reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
36.	Help students apply reading strategies to improve comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
37.	Use various types of assessments to	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

	motivate students to read?						
38.	Encourage uninterested students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
39.	Use a variety of assessments to help struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
40.	Help students choose which reading strategy to use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
41.	Increase students' confidence in reading?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
42.	Use a variety of texts to encourage students to read?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
43.	Select course readings to match the needs of individual students?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
44.	Communicate the importance of reading comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
45.	Adapt reading materials for struggling readers?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
46.	Identify improper reading strategy use?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
47.	Differentiate reading strategy instruction?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
48.	Help struggling readers self-assess their comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely
49.	Help struggling readers improve their comprehension?	Not at all	Very little	Somewhat	Moderately	Quite a bit	Completely

Please indicate your opinion about each of the statements below.

50. How much can you do to motivate students who show low interest in school work?

1 Nothing 2 3 Very little 4 5 Some influence 6 7 Quite a bit 8 9 A great deal

51. How well can you implement alternative strategies in your classroom?
1 Nothing 2 3 Very little 4 5 Some influence 6 7 Quite a bit 8 9 A great deal

52. To what extent can you craft good questions for your students?
1 Nothing 2 3 Very little 4 5 Some influence 6 7 Quite a bit 8 9 A great deal

53. How much can you do to help your students value learning?
1 Nothing 2 3 Very little 4 5 Some influence 6 7 Quite a bit 8 9 A great deal

54. How much can you use a variety of assessment strategies?
1 Nothing 2 3 Very little 4 5 Some influence 6 7 Quite a bit 8 9 A great deal

55. How much can you do to get students to believe they can do well in school work?
1 Nothing 2 3 Very little 4 5 Some influence 6 7 Quite a bit 8 9 A great deal
56. To what extent can you provide an alternative explanation or example when students are confused?
1 Nothing 2 3 Very little 4 5 Some influence 6 7 Quite a bit 8 9 A great deal

57. How much can you assist families in helping their children do well in school?
1 Nothing 2 3 Very little 4 5 Some influence 6 7 Quite a bit 8 9 A great deal

Please respond to how often these statements apply to you.

58. I am quick to recognize my own value as a teacher.
1 Never 2 3 4 5 6 7 8 9 Always

59. I like to include others in the activities in my classroom.
1 Never 2 3 4 5 6 7 8 9 Always

60. I plan time for enjoyable activities that I can do on my own.
1 Never 2 3 4 5 6 7 8 9 Always

61. I congratulate myself on my teaching successes.
1 Never 2 3 4 5 6 7 8 9 Always

62. I forgive my colleagues for their mistakes.
1 Never 2 3 4 5 6 7 8 9 Always

63. I am quick to recognize the value of others with whom I work.
1 Never 2 3 4 5 6 7 8 9 Always

64. I am impressed with my own teaching abilities.
1 Never 2 3 4 5 6 7 8 9 Always

65. I congratulate my colleagues on their successes.
1 Never 2 3 4 5 6 7 8 9 Always

66. I try not to be critical of the people with whom I work.
1 Never 2 3 4 5 6 7 8 9 Always

67. I pay attention to my own needs.
1 Never 2 3 4 5 6 7 8 9 Always

68. I am impressed with the abilities of others with whom I work.
1 Never 2 3 4 5 6 7 8 9 Always

69. I forgive myself for my mistakes in teaching.
1 Never 2 3 4 5 6 7 8 9 Always
Please rate your agreement with the following statements.

70. I'd feel I had a successful day in school if a student asked a question in class that made me think again about the subject matter.
1 Totally Disagree 2 3 4 5 Totally Agree

71. My relationships with students are more important to me than anything else in my role as teacher.
1 Totally Disagree 2 3 4 5 Totally Agree

72. My main goal as a teacher is to build a deep personal relationship with each and every student.
1 Totally Disagree 2 3 4 5 Totally Agree

73. I'd feel I had a successful day in school if I learned something new about myself as a teacher.
1 Totally Disagree 2 3 4 5 Totally Agree

74. I feel that I have succeeded as a teacher if I create close and warm relationships with students and classes.
1 Totally Disagree 2 3 4 5 Totally Agree

76. I'd feel I had a particularly successful day in school if I saw that I am developing professionally and teaching more effectively than in the past.
1 Totally Disagree 2 3 4 5 Totally Agree

76. I'd feel I had a successful day in school if something that happened in class made me want to deepen my professional knowledge.
1 Totally Disagree 2 3 4 5 Totally Agree

77. More than anything, I strive to create and maintain meaningful relationships with students.
1 Totally Disagree 2 3 4 5 Totally Agree

Thank you for completing the survey! If you would like to be entered into a drawing to win one of four $50 Target Gift Cards; please provide your email address below. Thank you! Reading Instruction Survey

78. If chosen, at what email address would you like the gift certificate to be sent?

Appendix W

Dr. Mraz Consent to Use Name in Dissertation

Dr. Mraz,

Good afternoon! My name is Erin Ramirez (Betty Sturtevant's student) and you had helped me last year as an expert reviewer on a scale I developed to assess secondary content-area teachers' sense of efficacy for reading instruction across the content areas.

First and foremost, thank you again for all of your assistance; I successfully defended my dissertation yesterday, and am very excited to move forward to confirm the factor structure of the instrument.

I'm reaching out to you because during defense my committee was concerned about me using my expert reviewers names without having explicit consent. Thus, what is your preference as far as being named within my dissertation? I can either use your name, or I can simply describe you without naming you specifically; I have no preference, and therefore wanted to check with you.

Thank you again so much for all of your help throughout this process; I greatly appreciate all of the assistance as I progressed through my doctorate.

Kind Regards, Erin Ramirez

Response:

Hello Erin,

Congratulations on completing your dissertation! And thank you for checking on my preference re: the name use. It's fine with me if you use my name in your dissertation.

Hope to meet you in person at a conference one of these days.

All the best,
Maryann

Appendix X

Dr. Woolfolk Hoy Consent to Use Name in Dissertation

Dr. Woolfolk Hoy,

Good afternoon! I hope all is well! I met you last year at the AEPC conference (I was the PowerPoint slide girl), and you helped me as an expert reviewer on a scale I developed to assess secondary content-area teachers' sense of efficacy for reading instruction across the content areas.

First and foremost, thank you again for all of your assistance; I successfully defended my dissertation yesterday, and am very excited to move forward to confirm the factor structure of the instrument.

I'm reaching out to you because during defense my committee was concerned about me using my expert reviewers names without having explicit consent. Thus, what is your preference as far as being named within my dissertation? I can either use your name, or I can simply describe you without naming you specifically; I have no preference, and therefore wanted to check with you.

Thank you again so much for all of your help throughout this process; I greatly appreciate all of the assistance as I progressed through my doctorate.

Kind Regards,

Erin Ramirez

Response:

I am happy to be listed by name if you like or not to be named, whatever others are doing.

Anita

Appendix Y

Dr. Nancy Frey Consent to Use Name in Dissertation

Dr. Frey,

Good afternoon! My name is Erin Ramirez (you were my mentor for the LRA DSICG) and you had helped me last year as an expert reviewer on a scale I developed to assess secondary content-area teachers' sense of efficacy for reading instruction across the content areas.

I'm reaching out to you because during defense my committee was concerned about me using my expert reviewers names without having explicit consent. Thus, what is your preference as far as being named within my dissertation? I can either use your name, or I can simply describe you without naming you specifically; I have no preference, and therefore wanted to check with you.

Thank you again so much for all of your help throughout this process; I greatly appreciate all of the assistance as I progressed through my doctorate.

Kind Regards,

Erin Ramirez

Response:

Hi Erin. I have no problem with being named, and whatever works for you is fine with me. Nancy

Appendix Z

Dr. Ellen Usher Consent to Use Name in Dissertation

Dr. Usher,

Good afternoon! I hope all is well! Last year you helped me as an expert reviewer on a scale I developed to assess secondary content-area teachers' sense of efficacy for reading instruction across the content areas.

I'm reaching out to you because during defense my committee was concerned about me using my expert reviewers names without having explicit consent. Thus, what is your preference as far as being named within my dissertation? I can either use your name, or I can simply describe you without naming you specifically; I have no preference, and therefore wanted to check with you.

Thank you again so much for all of your help throughout this process; I greatly appreciate all of the assistance as I progressed through my doctorate.

Kind Regards,

Erin Ramirez

Response:

Thanks for checking with me on the permissions. I'm fine with you mentioning my name. No problem at all.

Ellen

References

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61,* 364-373. doi: 10.1598/RT.61.5.1

Akbari, R., & Tavassoli, K. (2014). Developing an ELT context-specific teacher efficacy instrument. *RELC Journal*, *45*(1), 27-50. doi: 10.1177/0033688214523345

Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *17*(2), 86-95. doi: 10.1177/088840649401700203

Alvermann, D. (2001). *Effective literacy instruction for adolescents*. Executive summary and paper commissioned by the National Reading Conference. Chicago, IL: National Reading Conference.

Alvermann, D. E., Phelps, S., & Ridgeway, V. (2007). *Content area reading and literacy (5th Edition)*. Boston, MA: Allyn & Bacon.

Anderson, V., Chan, C. K. K., & Henne, R. (1995). The effects of strategy instruction on the literacy models of reading and writing delayed middle school students. In K. A. Hinchman, D. J. Leu, & C. K. Kinzer, (Eds.), *Perspectives on literacy research and practice: Forty-forth yearbook of the National Reading Conference*, pp. 180-189. Chicago, IL: National Reading Conference.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers*. Champaign, IL: University of Illinois. Center for the Study of Reading.

Armor, D., Conroy-Oseguera, P., Cox, M., King, N., McDonnell, L., & Pascal, A. P. E., & Zellman, G. (1976). *Analysis of the school preferred reading programs in selected Los Angeles minority schools.* (Report No. R-2007-LA USD). Santa Monica, CA: Rand Corporation.

Bain, R. (2006). Rounding up unusual suspects: Facing the authority hidden in the history classroom. *Teachers College Record*, *108*, 2080-2114. http://dx.doi.org/10.1111%2Fj.1467-9620.2006.00775.x

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. doi: 10.1037/0033-295X.84.2.191

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117-148. doi: 10.1207/s15326985ep2802_3

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Macmillan.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares, & T. C. Urdan, (Eds.) *Self-efficacy beliefs of adolescents* (pp. 307-337). Charlotte, NC: Information Age Publishing.

Barros, M. A., Laburú, C. E., & Silva, F. R. D. (2010). An instrument for measuring self-efficacy beliefs of secondary school physics teachers. *Procedia-Social and Behavioral Sciences, 2*, 3129-3133. http://dx.doi.org/10.1016/j.sbspro.2010.03.476

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, *3*(2), 77-85. doi: 10.1111/j.2044-8317.1950.tb00285.x

Bean, T. W. (2000). Reading in the content area: Social constructivist dimensions. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 629- 643). Mahwah, NJ: Lawrence Erlbaum Associates. http://dx.doi.org/10.4324/9781410605023

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, *71*, 287-311. doi: 10.1093/poq/nfm006

Behar-Horenstein, L. S., Pajares, F., & George, P. S. (1996). The effect of teachers' beliefs on students' academic performance during curriculum innovation. *The High School Journal*, *79,* 324-332.

Bell, B. S., & Kozlowski, W. J. (2002). Goal orientation and ability: interactive effects on self-efficacy, performance, and knowledge. *Journal of Applied Psychology*, *87*, 497-505.

Belson, W. (1981). *The design and understanding of survey questions*. Aldershot, England: Gower

Betoret, F. D. (2006). Stressors, self-efficacy, coping resources, and burnout among secondary school teachers in Spain. *Educational Psychology, 26*, 519-539. doi: 10.1080/01443410500342492

Biancarosa, G., & Snow, C. E. (2004). *Reading next: A vision for action and research in middle and high school literacy: A report from Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.

Biancarosa, G., & Snow, C. E. (2006). *Reading next: A vision for action and research in middle and high school literacy: A report to Carnegie corporation of New York (2nd ed.).* Washington, DC: Alliance for Excellent Education.

Bobbett, J. (2001). *School culture, teacher efficacy, and decision making in demonstrably effective and ineffective schools* (Doctoral Dissertation). Louisiana State University, Baton Rouge, Louisiana. Retrieved from ProQuest, UMI Dissertations Publishing. (Accession: 3021421).

Boulware-Gooden, R., Carreker, S., Thornhill, A., & Joshi, R. (2007). Instruction of metacognitive strategies enhances reading comprehension and vocabulary achievement of third-grade students. *The Reading Teacher*, *61*, 70-77. doi: 10.1598/RT.61.1.7

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, *16*, 239-253. http://dx.doi.org/10.1037/a0028613

Brown, R. (2002). Straddling two worlds: Self-directed comprehension instruction for middle schoolers. In C. C. Block, & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 337-350). New York, NY: Guilford Press.

Brozo, W. G., & Simpson, M. L. (2006). *Content literacy for today's adolescents: Honoring diversity and building competence*. Upper Saddle River, NJ: Pearson.

Bruce, C. D., Esmonde, I., Ross, J., Dookie, L., & Beatty, R. (2010). The effects of sustained classroom-embedded teacher professional learning on teacher efficacy and related student achievement. *Teaching and Teacher Education*, *26*, 1598-1608. doi:10.1016/j.tate.2010.06.011

Buehl, D. (2009). Mentoring literacy practices in academic disciplines. In S. R. Parris, D. Fisher, & K. Headley (Eds.), *Adolescent literacy, field tested: Effective solutions for every classroom* (pp. 228-238). Newark, DE: International Reading Association. http://dx.doi.org/10.1598/0695.19

Butler, R. (2007). Teachers' achievement goal orientations and associations with teachers' help seeking: Examination of a novel approach to teacher motivation. *Journal of Educational Psychology*, *99*, 241-252. http://dx.doi.org/10.1037/0022-0663.99.2.241

Butler, R. (2012). Striving to connect: Extending an achievement goal approach to teacher motivation to include relational goals for teaching. *Journal of Educational Psychology*, *104*, 726-742. http://dx.doi.org/10.1037/a0028613

Butler, R., & Shibaz, L. (2014). Striving to connect and striving to learn: Influences of relational and mastery goals for teaching on teacher behaviors and student interest and help seeking. *International Journal of Educational Research*, *65*, 41-53. doi:10.1016/j.ijer.2013.09.006

Cambria, J., & Guthrie, J. T. (2010). Motivating and engaging students in reading. *New England Reading Association Journal*, *46*(1), 16-29.

Campanelli, P. (1997). Testing survey questions: New directions in cognitive interviewing. *Bulletin de Methodologie Sociologique*, *55*, 5-17. doi: 10.1177/075910639705500103

Cantrell, S. C., & Callaway, P. (2008). High and low implementers of content literacy instruction: Portraits of teacher efficacy. *Teaching and Teacher Education*, *24*, 1739-1750. doi:10.1016/j.tate.2008.02.020

Cantrell, S. C., & Hughes, H. K. (2008). Teacher efficacy and content literacy implementation: An exploration of the effects of extended professional development with coaching. *Journal of Literacy Research*, *40*, 95-127. doi: 10.1080/10862960802070442

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology*, *95*, 821-832. http://dx.doi.org/10.1037/0022-0663.95.4.821

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*, 473-490. http://dx.doi.org/10.1016/j.jsp.2006.09.001

Carroll, J. B. (1953). An analytical solution for approximating simple structure in factor analysis. *Pyschometrika*, *18*(1), 23-28. doi: 10.1007/BF02289025

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245-276. doi: 10.1207/s15327906mbr0102_10

Chacón, C. T. (2005). Teachers' perceived efficacy among English as a foreign language teachers in middle schools in Venezuela. *Teaching and Teacher Education*, *21*, 257-272. doi:10.1016/j.tate.2005.01.001

Chan, D. W. (2006). Emotional intelligence and components of burnout among Chinese secondary school teachers in Hong Kong. *Teaching and Teacher Education*, *22*, 1042-1054. doi:10.1016/j.tate.2006.04.005

Clark, K. F., & Graves, M. F. (2005). Scaffolding students' comprehension of text. *The Reading Teacher*, *58*, 570-580. doi: 10.1598/RT.58.6.6

Cliff, N. (1970). The relation between sample and population characteristic vectors. *Pyschometrika, 35,* 163-178. doi: 10.1007/BF02291261

Coladarci, T., & Fink, D. R. (1995, April). *Correlations among measures of teacher efficacy: Are they measuring the same thing*. At the annual meeting of the American Educational Research Association, San Francisco, CA.

Comfrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology, 56*, 754-761. doi: 10.1037/0022-006x.56.5.754

Conley, M. W. (2009). Improving adolescent comprehension: Developing comprehension strategies in the content areas. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 531-550). New York, NY: Routledge.

Conley, M. W., & Hinchman, K. A. (2004). No child left behind: What it means for US adolescents and what we can do about it. *Journal of Adolescent & Adult Literacy*, *48*, 42-50. doi: 10.1598/JAAL.48.1.4

Conrad, F., & Blair, J. (2004). Data quality in cognitive interviews: the case of verbal reports. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 67-87). Hoboken, NJ: John Wiley & Sons.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104. doi: 10.1037/0021-9010.78.1.98

Cotterall, S., & Murray, G. (2009). Enhancing metacognitive knowledge: Structure, affordances and self. *System*, *37*(1), 34-45. http://dx.doi.org/10.1016/j.system.2008.08.003

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando, FL: Holt, Rinehart and Winston.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334. doi: 10.1007/BF02310555

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, *16*, 137-163. doi: 10.1111/j.2044-8317.1963.tb00206.x

Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology*, *80*(2), 131-142. doi: http://dx.doi.org/10.1037/0022-0663.80.2.131

Cureton, E. E., & D'Agostino, R. B. (1983). *Factor analysis: An applied approach.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching.* New York, NY: National Commission on Teaching & America's Future.

Deemer, S. (2004). Classroom goal orientation in high school classrooms: Revealing links between teacher beliefs and classroom environments. *Educational Research*, *46*, 73-90. doi:10.1080/0013188042000178836

Dellinger, A. B. (2001). *A study of the measurement and sources of teachers' self and collective efficacy beliefs in professional learning environments* (Doctoral Dissertation). Louisiana State University, Baton Rouge, Louisiana. Retrieved from ProQuest, UMI Dissertations Publishing. (Accession: 3021424).

Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education, 24*, 751-766. http://dx.doi.org/10.1016/j.tate.2007.02.010

De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, *35*, 174-192. doi:10.1016/j.cedpsych.2010.03.001

DeMaio, T. J., & Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz and S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research* (pp. 177-195). San Francisco, CA: Jossey-Bass.

Denton, C. A., & Fletcher, J.M. (2003). Scaling reading interventions. In B.R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bring science to scale* (pp.445-464). Timonium, MD: York Press.

DeVellis R.F. (2003). *Scale development: Theory and applications*. Thousand Oaks, CA: SAGE Publications.

Di Fabio, A., & Palazzeschi, L. (2008). Emotional intelligence and self-efficacy in a sample of Italian high school teachers. *Social Behavior & Personality: An International Journal*, *36*, 315-326. http://dx.doi.org/10.2224/sbp.2008.36.3.315

Dimitrov, D. M. (2012). *Statistical methods for validation of assessment scale data in counseling and related fields*. Alexandria, VA: American Counseling Association.

Duffy, G. G. (2002). The case for direct explanation of strategies. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 28-41). New York, NY: Guilford Press.

Duffin, L. C., French, B. F., & Patrick, H. (2012). The teachers' sense of efficacy scale: Confirming the factor structure with beginning pre-service teachers. *Teaching and Teacher Education*, *28*, 827-834. http://dx.doi.org/10.1016/j.tate.2012.03.004

Duke, N. K. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly*, *35*, 202-224. doi: 10.1598/RRQ.35.2.1

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. *What research has to say about reading instruction* (3rd ed.), (pp. 205-242). Newark, DE: International Reading Association. doi: 10.1598/0872071774.10

Ebrahimi, M. R., & Moafian, F. (2012). Does emotional intelligence or self-efficacy have something to do with high school English teachers' critical thinking, considering demographic Information? *International Journal of Linguistics*, *4,* 224-242. doi: http://dx.doi.org/10.5296/ijl.v4i4.2662

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*, 262-300. doi: 10.3102/0034654308325998

Elkins, J., & Luke, A. (1999). Redefining adolescent literacies. *Journal of Adolescent & Adult Literacy*, *43*, 212-215.

Enochs, L. G., Riggs, I. M., & Ellis, J. D. (1993). The development and partial validation of microcomputer utilization in teaching efficacy beliefs instrument in a science setting. *School Science and Mathematics*, *93*, 257-263. doi: 10.1111/j.1949-8594.1993.tb12240.x

Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the mathematics teaching efficacy beliefs instrument. *School Science and Mathematics, 100*, 194-202. doi: 10.1111/j.1949-8594.2000.tb17256.x

Everitt, B. S. (1975). Multivariate analysis: The need for data, and other problems. *The British Journal of Psychiatry*, *126*, 237-240. doi:10.1192/bjp.126.3.237

Evertson, C.M. & Green, J.L. (1986). Observation as inquiry and method. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 162-213). New York: Macmillan.

Evers, W. J., Brouwers, A., & Tomic, W. (2002). Burnout and self-efficacy: A study on teachers' beliefs when implementing an innovative educational system in the Netherlands. *British Journal of Educational Psychology*, *72*, 227-243. doi: 10.1348/000709902158865

Fisher, D., & Frey, N. (2008). What does it take to create skilled readers? Facilitating the transfer and application of literacy strategies. *Voices from the Middle, 15,* 416-22. doi: 10.1080/19388070802300330

Fisher, D., & Frey, N. (2011). *Improving adolescent literacy: Content area strategies at work*. New York, NY: Pearson Higher Ed.

Fisher, D., Frey, N., & Nelson, J. (2012). Literacy achievement through sustained professional development. *The Reading Teacher*, *65*, 551-563. doi: 10.1002/TRTR.01082

Fisher, D., Frey, N., & Williams, D. (2002). Seven literacy strategies that work. *Educational Leadership*, *60*(3), 70-73.

Fisher, D., & Ivey, G. (2005). Literacy and language as learning in content-area classes: A departure from "Every teacher a teacher of reading." *Action in Teacher Education*, *27*(2), 3-11. doi: 10.1080/01626620.2005.10463378

Fives, H., & Buehl, M. M. (2010). Examining the factor structure of the teachers' sense of efficacy scale. *The Journal of Experimental Education, 78*, 118-134. doi: 10.1080/00220970903224461

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Lawrence Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist*, 34, 906-911. doi: http://dx.doi.org/10.1037/0003-066x.34.10.906

Freedman, L., & Carver, C. (2007). Pre-service teacher understandings of adolescent literacy development: Naïve wonder to dawning realization to intellectual rigor. *Journal of Adolescent & Adult Literacy*, *50*, 654-665. doi: 10.1598/JAAL.50.8.4

Froman, R. D. (2001). Elements to consider in planning the use of factor analysis. *Southern Online Journal of Nursing Research*, *5*(2), 1-22. Retrieved November 6, 2014 from http://www.snrs.org/ members/SOJNR_articles/iss05vol02.pdf

Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex Publishing.

Gee, T. C, & Forrester, N. (1988). Moving reading instruction beyond the reading classroom. *Journal of Reading, 31*, 505-510.

Gehlbach, H., & Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology*, *15*, 380-387. doi: http://dx.doi.org/10.1037/a0025704

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, *71*, 279-320. doi: http://dx.doi.org/10.3102/00346543071002279

Ghaith, G., & Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 13*, 451-458. doi:10.1016/S0742-051X(96)00045-5

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*, 569-582. doi.:10.1037/0022-0663.76.4.569

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17,* 183-211. doi: 10.2307/258770

Gooden, S. H. (2012). Comprehension strategies teachers use when they read. *Journal of Reading Education*, *37*(2), 16-20.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gough, P. B., & Hillinger, M. L. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society*, *30*, 179-196. doi: http://dx.doi.org/10.1007/bf02653717

Gray, W. S. (1925). *Summary of investigations relating to reading.* (Supplementary Educational Monographs, No. 28). Chicago, IL: University of Chicago Press.

Greenleaf, C. L., Schoenbach, R., Cziko, C., & Mueller, F. L. (2001). Apprenticing adolescent readers to academic literacy. *Harvard Educational Review*, *71*, 79-130. doi: http://dx.doi.org/10.17763/haer.71.1.q811712577334038

Grigg, W. S., Donahue, P., & Dion, G. (2007). *The nation's report card: 12th grade reading and mathematics 2005* (NCES 2007-468). Washington DC: U.S. Department of Education, National Center for Educational Statistics.

Guo, Y., Connor, C. M., Yang, Y., Roehrig, A. D., & Morrison, F. J. (2012). The effects of teacher qualification, teacher self-efficacy, and classroom practices on fifth graders' literacy outcomes. *The Elementary School Journal*, *113*, 3-24. doi: 10.1086/665816

Guo, Y., Dynia, J. M., Pelatti, C. Y., & Justice, L. M. (2014). Self-efficacy of early childhood special education teachers: Links to classroom quality and children's learning for children with language impairment. *Teaching and Teacher Education*, *39*, 12-21. doi:10.1016/j.tate.2013.11.005

Guo, Y., Justice, L. M., Sawyer, B., & Tompkins, V. (2011). Exploring factors related to preschool teachers' self-efficacy. *Teaching and Teacher Education*, *27*, 961-968. http://dx.doi.org/10.1016/j.tate.2011.03.008

Guo, Y., Piasta, S. B., Justice, L. M., & Kaderavek, J. N. (2010). Relations among preschool teachers' self-efficacy, classroom quality, and children's language and literacy gains. *Teaching and Teacher Education, 26,* 1094-1103. http://dx.doi.org/10.1016/j.tate.2009.11.005

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. doi: 10.1177/002248718103200310

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, *4*, 63-69. doi:10.1016/0742-051X(88)90025-X

Guskey, T. R. (2000). *Evaluating professional development and learning*. Newbury Park, CA: Corwin Press.

Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal, 31*, 627-643. doi: 10.3102/00028312031003627

Guthrie, J. T., & Wigfield, A. (1999). How motivation fits into a science of reading. *Scientific Studies of Reading*, *3*, 199-207. doi: 10.1207/s1532799xssr0303_1

Guthrie, J. T., & Wigfield, A. (2000) Engagement and motivation in reading. In M. L. Kamil & P. B. Mosenthal (Eds.), *Handbook of reading research, Vol. III*. (pp. 403-422). Mahwah, NJ: Erlbaum.

Guthrie, J. T., Van Meter, P., McCann, A., Wigfield, A., Bennett, L., Poundstone, C., Rice, M. E., Fabisch, F., Hunt, B., & Mitchell, A. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly*, *31*, 306-332. doi: 10.1598/RRQ.31.3.5

Guthrie, J. T., Wigfield, A., & Perencevich, K. C. (2004). Scaffolding for motivation and engagement in reading. *Motivating reading comprehension: Concept-Oriented Reading Instruction.* Mahwah, NJ: Lawrence Erlbaum. doi: 10.4324/9781410610126

Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology*, *92*, 331-341. http://dx.doi.org/10.1037/0022-0663.92.2.331

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M., Scaffidi, N., & Tonks, S. (2004). Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. *Journal of Educational Psychology, 96*(3), 1-21. doi: 10.1037/0022-0663.96.3.403

Hall, L. A. (2005). Teachers and content area reading: Attitudes, beliefs, and change. *Teaching and Teacher Education, 21*, 403-414. http://dx.doi.org/10.1016/j.tate.2005.01.009

Haverback, H. R. (2007). *Pre-service reading teachers as tutors: An examination of efficacy and content knowledge* (Doctoral dissertation, Doctoral dissertation, University of Maryland, College Park). Retrieved April 24, 2013 from http://search.proquest.com.mutex.gmu.edu/docview/304855008?accountid=14541

Haverback, H. R., & Parault, S. J. (2011). High efficacy and the pre-service reading teacher: A comparative study. *Teaching and Teacher Education*, *27*, 703-711. http://dx.doi.org/10.1016/j.tate.2010.12.001

Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement*. Washington, DC: Alliance for Excellent Education.

Henson, R. K. (2002). From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist, 37*, 137-150. doi: 10.1207/S15326985EP3703_1

Henson, R. K., Kogan, L. R., & Vacha-Haase, T. (2001). A reliability generalization study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement, 61*, 404-420. doi: 10.1177/00131640121971284

Herber, H. L. (1978). *Teaching reading in content areas*. Englewood Cliffs, NJ: Prentice-Hall.

Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to instructional quality: A longitudinal analysis. *Journal of Educational Psychology, 105*, 774-786. http://dx.doi.org/10.1037/a0032198

Holzberger, D., Philipp, A., & Kunter, M. (2014). Predicting teachers' instructional behaviors: The interplay between self-efficacy and intrinsic needs. *Contemporary Educational Psychology, 39*, 100-111. doi:10.1016/j.cedpsych.2014.02.001

Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education, 21*, 343-356. http://dx.doi.org/10.1016/j.tate.2005.01.007

Hoy, W. K., & Woolfolk, A. E. (1993). Teachers' sense of efficacy and the organizational health of schools. *The Elementary School Journal*, *93*, 355-372. doi: 10.1086/461729

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

International Reading Association. (2012). *Adolescent literacy* (Position statement, Rev. 2012 ed.). Newark, DE: International Reading Association.

Irvin, J. L., Meltzer, J., & Dukes, M. (2007). *Taking action on adolescent literacy: An implementation guide for school leaders.* Alexandria, VA: ASCD.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141-151. doi: http://dx.doi.org/10.1177/001316446002000116

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. doi: 10.1007/BF02291575

Kamil, M. L. (2003). *Adolescents and literacy: Reading for the 21st century*. Washington, DC: Alliance for Excellent Education.

Kamil, M. L., & Bernhardt, E. B. (2009). Introduction. In E. B. Huey, *The psychology and pedagogy of reading*. Newark, DE: International Reading Association.

Karabenick, S. A., Woolley, M. E., Friedel, J. M., Ammon, B. V., Blazevski, J., Bonney, C. R., & Kelly, K. L. (2007). Cognitive processing of self-report items in educational research: Do they think what we mean? *Educational Psychologist*, *42*, 139-151. doi:10.1080/00461520701416231

Kerlinger, F. N. (1973). *Foundations of behavioral research* (2nd ed.). New York, NY: Holt, Rinehart and Winston.

Kim, Y. H., & Goetz, E. T. (1994). Context effects on word recognition and reading comprehension of poor and good readers: A test of the interactive-compensatory hypothesis. *Reading Research Quarterly*, *29*, 179-188. doi: 10.2307/747810

Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y., & Georgiou, T. (2009). Exploring the validity of a teachers’ self-efficacy scale in five countries. *Contemporary Educational Psychology*, *34*, 67-76. http://dx.doi.org/10.1016/j.cedpsych.2008.08.001

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology, 102*, 741-756. http://dx.doi.org/10.1037/a0019237

Klassen, R. M., Tze, V. M., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998–2009: signs of progress or unfulfilled promise? *Educational Psychology Review, 23,* 21-43. doi: 10.1007/s10648-010-9141-8

Klassen, R. M., Yerdelen, S., & Durksen, T. L. (2013). Measuring teacher engagement: Development of the engaged teachers scale (ETS). *Frontline Learning Research*, *1*(2), 33-52. http://dx.doi.org/10.14786/flr.v1i2.44

Kline, R. B. (2011). *Principles and practice of structural equation modeling (3rd Edition)*. New York, NY: Guilford Press.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, *50*(1), 537-567. doi: 10.1146/annurev.psych.50.1.537

Kurz, T. B., & Knight, S. L. (2004). An exploration of the relationship among teacher efficacy, collective teacher efficacy, and goal consensus. *Learning Environments Research*, *7,* 111-128. doi: 10.1023/B:LERI.0000037198.37750.0e

Lee, V. E., Dedrick, R. F., & Smith, J. B. (1991). The effect of the social organization of schools on teachers' efficacy and satisfaction. *Sociology of Education*, *64,* 190-208. doi: 10.2307/2112851

Lee, C. D., & Spratley, A. (2010). *Reading in the disciplines: The challenges of adolescent literacy*. New York, NY: Carnegie Corporation of New York.

Luke, A., & Elkins, J. (2000). Special themed issue: Re/mediating adolescent literacies. *Journal of Adolescent & Adult Literacy*, *43*, 396-398.

Martin, N. K., Sass, D. A., & Schmitt, T. A. (2012). Teacher efficacy in student engagement, instructional management, student stressors, and burnout: a theoretical model using in-class variables to predict teachers' intent-to-leave. *Teaching and Teacher Education*, *28,* 546-559. doi:10.1016/j.tate.2011.12.003

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works* (Vol. 5). Alexandria, VA: Association for Supervision and Curriculum Development.

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach: An interactive approach*. Thousand Oaks, CA: Sage.

May, S. (2007). Sustaining effective literacy practices over time in secondary schools: School organisational and change issues. *Language and Education*, *21*, 387-405. doi: 10.2167/le799.0

McConachie, S. M., & Petrosky, A. R. (2009). *Content matters: A disciplinary literacy approach to improving student learning*. Hoboken, NJ: John Wiley & Sons. http://dx.doi.org/10.1002/9781118269466

McNeill, K. L., Pimentel, D. S., & Strauss, E. G. (2013). The impact of high school science teachers' beliefs, curricular enactments and experience on student learning during an inquiry-based urban ecology curriculum. *International Journal of Science Education*, *35*, 2608-2644. doi:10.1080/09500693.2011.618193

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97. http://dx.doi.org/10.1037/h0043158

Moje, E. (2000). "To be part of the story": The literacy practices of gangsta adolescents. *Teachers College Record*, *102*, 651-690. doi: 10.1111/0161-4681.00071

Moje, E. B. (2007). Developing socially just subject-matter instruction: A review of the literature on disciplinary literacy teaching. *Review of Research in Education*, *31*(1), 1-44. doi: 10.3102/0091732X07300046001

Moje, E. B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent & Adult Literacy*, *52*, 96-107. doi: 10.1598/JAAL.52.2.1

Moje, E. B., Collazo, T., Carrillo, R., & Marx, R. W. (2001). "Maestro, what is 'quality'?": Language, literacy, and discourse in project-based science. *Journal of Research in Science Teaching*, *38*, 469-498. doi: 10.1002/tea.1014

Moje, E. B., Overby, M., Tysvaer, N., & Morris, K. (2008). The complex world of adolescent literacy. *Harvard Educational Review*, *78*, 107-154. doi: 10.1598/JAAL.52.2.1

Moje, E. B., Young, J. P., Readence, J. E., & Moore, D. W. (2000). Reinventing adolescent literacy for new times: Perennial and millennial issues. *Journal of Adolescent and Adult Literacy*, *43*, 400-411. doi: 10.1080/19388070209558367

Moje, E. B. & Speyer, J. (2008). The reality of challenging texts in high school science and social studies: How teachers can mediate comprehension. In K. Hinchman, & H. Thomas (Eds.), *Best practices in adolescent literacy instruction* (pp. 185-211). New York, NY: Guilford.

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94*, 249-259. doi: 10.1037/0022-0663.94.2.249

Moore, D. W., Readence, J. E., & Rickelman, R. J. (1983). An historical exploration of content area reading instruction. *Reading Research Quarterly*, *18*, 419-438. doi: 10.1080/19388078409557780

Monte-Sano, C. (2011). Beyond reading comprehension and summary: Learning to read and write in history by focusing on evidence, perspective, and interpretation. *Curriculum Inquiry*, *41*, 212-249. doi: 10.1111/j.1467-873X.2011.00547.x

Monte-Sano, C., & De La Paz, S. (2012). Using writing tasks to elicit adolescents' historical reasoning. *Journal of Literacy Research, 44*, 273-299. doi: 10.1177/1086296X12450445

Monte-Sano, C., De La Paz, S., & Felton, M. (2014). Implementing a disciplinary-literacy curriculum for US history: learning from expert middle school teachers in diverse classrooms. *Journal of Curriculum Studies*, *46*(4), 1-36. doi:10.1080/00220272.2014.904444

Moore, D., Zabrucky, K., & Commander, N. E. (1997). Metacomprehension and comprehension performance in younger and older adults. *Educational Gerontology: An International Quarterly*, *23*, 467-475. doi: http://dx.doi.org/10.1080/0360126970230506

Morris, D. (2010). *Sources of teaching self-efficacy: A scale validation* (Doctoral Dissertation). Emory University, Atlanta, GA. Retrieved from ProQuest, UMI Dissertations Publishing. (Accession: 3423100).

Moss, B. (2005). Making a case and a place for effective content-area literacy instruction and knowledge in the elementary grades. *The Reading Teacher*, *59,* 46-55. doi: 10.1598/RT.59.1.5

National Center for Education Statistics (2013). The Nation's Report Card: A First Look: 2013 Mathematics and Reading (NCES 2014-451). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Council of Teachers of English. (2006). *NCTE principals of adolescent literacy reform: A policy research brief.* Retrieved December 11, 2013 from http://www.ncte.org/library/NCTEFiles/resources/Positions/Adol-Lit-Brief.pdf

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

Ness, M. (2007). Reading comprehension strategies in secondary content-area classrooms. *Phi Delta Kappan, 89*, 229-232. doi: 10.1177/003172170708900314

Nokes, J. D., and Dole, J. A. (2004). Helping adolescent readers through explicit strategy instruction. In Jetton, T. L. and Dole, J. A. (Eds.), *Adolescent literacy research and practice*, (pp. 162-182). New York, NY: The Guilford Press.

O'Brien, D. G., Stewart, R. A., & Moje, E. B. (1995). Why content literacy is difficult to infuse into the secondary school: Complexities of curriculum, pedagogy, and school culture. *Reading Research Quarterly*, *30*, 442-463. doi: 10.2307/747625

Ofodu, G. O., & Adedipe, T. H. (2011). Assessing ESL students' awareness and application of metacognitive strategies in comprehending academic materials. *Journal of Emerging Trends in Educational Research and Policy Studies*, *2*, 343-346.

Olivier, D. (2000). *Teacher personal and school culture characteristics in effective schools: Toward a model of a professional learning community* (Doctoral Dissertation). Baton Rouge, LA: Louisiana State University. Retrieved from ProQuest, UMI Dissertations Publishing. (Accession: 3016566).

Olson, M. R., & Truxaw, M. P. (2009). Pre-service science and mathematics teachers and discursive metaknowledge of text. *Journal of Adolescent & Adult Literacy*, *52*, 422-431. doi: 10.1598/JAAL.52.5.6

Osborne, J. W. (2014). *Best practices in exploratory factor analysis*. Louisville, NY: CreateSpace Independent Publishing Platform.

Osborne, J. W., & Costello, A. B. (2005). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research & Evaluation*, *9*(11). Retrieved February 2, 2014 from http://PAREonline.net/getvn.asp?v=9&n=11

Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*, 307-332. doi: 10.3102/00346543062003307

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66,* 543-578. doi: 10.3102/00346543066004543

Pajares, F. (1997). Current directions in self-efficacy research. In Maehr, M. & Pintrich, P.R. (Eds.) *Advances in motivation and achievement* (Vol. 10., pp. 1-49). Greenwich, CT: JAI Press.

Pajares, F., & Johnson, M. J. (1996). Self-efficacy beliefs and the writing performance of entering high school students. *Psychology in the Schools*, *33*, 163-175. doi: 10.1002/(SICI)1520-6807(199604)33:2<163::AID-PITS10>3.0.CO;2-C

Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30*, 111-140. doi: 10.3102/0162373708317680

Paneque, O. M., & Barbetta, P. M. (2006). A study of teacher efficacy of special education teachers of English language learners with disabilities. *Bilingual Research Journal, 30,* 171-193. doi: 10.1080/15235882.2006.10162871

Pardo, L. S. (2004). What every teacher needs to know about comprehension. *The Reading Teacher*, *58*, 272-280. doi: 10.1598/RT.58.3.5

Paris, S. G., & Lindauer, B. K. (1982). The development of cognitive skills during childhood. In B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 333-349). Englewood Cliffs, NJ: Prentice-Hall.

Paris, S. G., & Oka, E. R. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review*, *6*, 25-56. http://dx.doi.org/10.1016/0273-2297(86)90002-x

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: SAGE Publications.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Pearson, P. D., & Dole, J. A. (1987). Explicit comprehension instruction: A review of research and a new conceptualization of instruction. *The Elementary School Journal*, *88*, 151-165. doi: 10.1086/461530

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology, 8*, 317-344. http://dx.doi.org/10.1016/0361-476X(83)90019-X

Perfetti, C. A., & Landi, N., Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (227-247). Malden, MA: Blackwell Publishing.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40. http://dx.doi.org/10.1037//0022-0663.82.1.33

Pressley, M. (2002). Metacognition and self-regulated comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research had to say about reading instruction* (3rd ed., pp. 291-309). Newark, DE: International Reading Association. doi: 10.1598/0872071774.13

Pressley, M. (2006). *Reading instruction that works: The case for balanced teaching*. New York, NY: Guilford Press.

Pressley, M., & Gaskins, I. W. (2006). Metacognitively competent reading comprehension is constructively responsive reading: How can such reading be developed in students? *Metacognition and Learning, 1*, 99-113. doi: 10.1007/s11409-006-7263-7

Pressley, M., & Harris, K. R. (1990). What we really know about strategy instruction. *Educational Leadership, 48*(1), 31-34.

Pressley, M., Wharton-McDonald, R., Hampston, J.M., & Echevarria, M. (1998). The nature of literacy instruction in ten grade-4 and -5 classrooms in upstate New York. *Scientific Studies of Reading*, 2, 150-191. doi: http://dx.doi.org/10.1207/s1532799xssr0202_4

Purkey, W. W. (2000). *What students say to themselves: Internal dialogue and school success*. Thousand Oaks, CA: Corwin Press.

Purkey, W. W., & Novak, J. M. (1996). *Inviting school success: A self-concept approach to teaching, learning, and democratic practice*. Florence, KY: Wadsworth, Inc.

RAND (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: Rand Corporation.

Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education*, *65*, 150-167. doi: 10.2307/2112680

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, *12*, 287–297. doi: http://dx.doi.org/10.1037/1040-3590.12.3.287

Retelsdorf, J., Butler, R., Streblow, L., & Schiefele, U. (2010). Teachers' goal orientations for teaching: Associations with instructional practices, interest in teaching, and burnout. *Learning and Instruction*, *20*(1), 30-46. doi:10.1016/j.learninstruc.2009.01.001

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education, 74*, 625-637. doi: 10.1002/sce.3730740605

Ritter, J. M., Boone, W. J., & Rubba, P. A. (2001). Development of an instrument to assess prospective elementary teacher self-efficacy beliefs about equitable science teaching and learning (SEBEST). *Journal of Science Teacher Education*, *12*, 175-198. http://dx.doi.org/10.1023/a:1016747713585

Rose, J. S., & Medway, F. J. (1981). Measurement of teachers' beliefs in their control over student outcome. *The Journal of Educational Research, 74*, 185-190. doi:10.1080/00220671.1981.10885308

Ross, J. A. (1992). Teacher efficacy and the effects of coaching on student achievement. *Canadian Journal of Education/Revue Canadienne de L'education, 17*(1), 51-65. doi: 10.2307/1495395

Ross, J. A., Cousins, J., & Gadalla, T. (1996). Within-teacher predictors of teacher efficacy. *Teaching and Teacher Education, 12*, 385-400. http://dx.doi.org/10.1016/0742-051x(95)00046-m

Rotter, J. B. (1966). Generalized expectancies' for internal versus external control of reinforcement. *Psychological Monographs, 80*(1), 1-28. doi: 10.1037/h0092976

Rubin, A., & Babbie, E. (2007). *Research methods for social work*. Belmont, CA: Thompson Brooks.

Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*, 57-67. doi: 10.1023/A:1007999204543

Shanahan, T. & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review, 78*, 40-61. doi: http://dx.doi.org/10.17763/haer.78.1.v62444321p602101

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*, 611-625. doi: 10.1037/0022-0663.99.3.611

Slater, W. H., & Horstman, F. R. (2002). Teaching reading and writing to struggling middle school and high school students: The case for reciprocal teaching. *Preventing School Failure: Alternative Education for Children and Youth*, *46*, 163-166. doi: 10.1080/10459880209604416

Smith, E. V., Wakely, M. B., De Kruif, R. E., & Swartz, C. W. (2003). Optimizing rating scales for self-efficacy (and other) research. *Educational and Psychological Measurement*, *63*, 369-391. doi: 10.1177/0013164403063003002

Smolleck, L. D., Zembal-Saul, C., & Yoder, E. P. (2006). The development and validation of an instrument to measure pre-service teachers' self-efficacy in regard to the teaching of science as inquiry. *Journal of Science Teacher Education, 17,* 137-163. doi: 10.1007/s10972-006-9015-6

Snow, C., & Moje, E. (2010). Why is everyone talking about adolescent literacy? *Phi Delta Kappan*, *91*(6), 66-69. doi: 10.1177/003172171009100616

Soodak, L. C., & Podell, D. M. (1996). Teacher efficacy: Toward the understanding of a multi-faceted construct. *Teaching and Teacher Education, 12,* 401-411. http://dx.doi.org/10.1016/0742-051x(95)00047-n

Stevens, T., Olivárez, A., & Hamman, D. (2006). The role of cognition, motivation, and emotion in explaining the mathematics achievement gap between Hispanic and White students. *Hispanic Journal of Behavioral Sciences*, *28*, 161-186. doi: 10.1177/0739986305286103

Sturtevant, E. G. (1996). Lifetime influences on the literacy-related instructional beliefs of experienced high school history teachers: Two comparative case studies. *Journal of Literacy Research*, *28*, 227-257. doi: 10.1080/10862969609547920

Sturtevant, E. G., & Linek, W. M. (2003). The instructional beliefs and decisions of middle and secondary teachers who successfully blend literacy and content. *Literacy Research and Instruction*, *43,* 74-89. doi: 10.1080/19388070309558401

Switzer, G. E., Wisniewski, S. R., Belle, S. H., Dew, M. A., & Schultz, R. (1999). Selecting, developing, and evaluating research instruments. *Social Psychiatry and Psychiatric Epidemiology*, *34*, 399-409. doi: 10.1007/s001270050161

Szabo, S. M., & Mokhtari, K. (2004). Developing a reading teaching efficacy instrument for teacher candidates: A validation study. *Action in Teacher Education*, *26*, 59-72. doi: 10.1080/01626620.2004.10463333

Taboada, A., & Guthrie, J. T. (2006). Contributions of student questioning and prior knowledge to construction of knowledge from reading information text. *Journal of Literacy Research*, *38*, 1-35. doi: 10.1207/s15548430jlr3801_1

Taboada, A., & Rutherford, V. (2011). Developing reading comprehension and academic vocabulary for English language learners through science content: A formative experiment. *Reading Psychology*, *32*, 113-157. doi: 10.1080/02702711003604468

Tang, H. N., & Moore, D. W. (1992). Effects of cognitive and metacognitive pre-reading activities on the reading comprehension of ESL Learners. *Educational Psychology*, *12*, 315-331. doi: 10.1080/0144341920120313

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

Thorndike. E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *The Journal of Educational Psychology*, *8*, 323-332. doi: 10.1037/h0075325

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago, IL: University of Chicago Press.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, United Kingdom: Cambridge University Press.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*, 783-805. http://dx.doi.org/10.1016/s0742-051x(01)00036-1

Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and teacher Education*, *23*, 944-956. doi:10.1016/j.tate.2006.05.003

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202-248. doi: 10.3102/00346543068002202

Tschannen-Moran, M., & Johnson, D. (2011). Exploring literacy teachers' self-efficacy beliefs: Potential sources at play. *Teaching and Teacher Education, 27*, 751-761. http://dx.doi.org/10.1016/j.tate.2010.12.005

Underwood, T., & Pearson, P. D. (2004). Teaching struggling adolescent readers to comprehend what they read. In Jetton, T. L. and Dole, J. A. (Eds.), *Adolescent literacy research and practice*, (pp. 135-161). New York, NY: The Guilford Press.

Usher, E. L., & Pajares, F. (2006). Inviting confidence in school: Invitations as a critical source of the academic self-efficacy beliefs of entering middle school students. *Journal of Invitational Theory and Practice*, *12*, 7-16.

U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS), "Public School Teacher and Private School Teacher Data Files," 2011–12.

Vacca, R. T. (1998). Let's not marginalize adolescent literacy. *Journal of Adolescent & Adult Literacy*, *41*, 604-610. http://dx.doi.org/10.1598/jaal.41.8.1

Vacca, R. T. (2002). From efficient decoders to strategic readers. *Educational Leadership*, *60*(3), 6-11.

Valiante, G., & Pajares, F. (1999). Invitations, motivation, and academic achievement. *Journal of Invitational Theory and Practice, 6*, 28-47.

van Uden, J. M., Ritzen, H., & Pieters, J. M. (2013). I think I can engage my students. Teachers' perceptions of student engagement and their beliefs about being a teacher. *Teaching and Teacher Education*, *32*, 43-54. doi:10.1016/j.tate.2013.01.004

Vaughn, S., Wanzek, J., Murray, C. S., Scammacca, N., Linan-Thompson, S., & Woodruff, A. L. (2009). Response to early reading intervention examining higher and lower responders. *Exceptional Children*, *75*, 165-183. doi: 10.1007/978-0-387-49053-3_17

Warren, L. L., & Payne, B. D. (1997). Impact of middle grades' organization on teacher efficacy and environmental perceptions. *The Journal of Educational Research*, *90*, 301-308. http://dx.doi.org/10.1080/00220671.1997.10544586

Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education*, *21*, 747-766. http://dx.doi.org/10.1016/j.tate.2005.05.009

Wigfield, A., Guthrie, J. T., Tonks, S., & Perencevich, K. C. (2004). Children's motivation for reading: Domain specificity and instructional influences. *The Journal of Educational Research*, *97*, 299-310. doi:10.3200/JOER.97.6.299-310

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.

Wineburg, S.S. (1991). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal*, *28*, 495-519. doi:10.2307/1163146.

Wood, K.D. (2003). New dimensions in content area literacy—not just for secondary teachers. *California Reader*, *36*(1), 12–17.

Woolfolk, A. E., & Hoy, W. K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology*, *82*, 81-91. http://dx.doi.org/10.1037/0022-0663.82.1.81

Worthy, J., & Invernizzi, M. A. (1995). Linking reading with meaning: A case study of a hyperlexic reader. *Journal of Literacy Research*, *27*, 585-603. doi: 10.1080/10862969509547900

Wren, S. (2002). Ten myths of reading instruction. *SEDL Letter*, *14*(3), 3-8.

Yang, N. D. (1999). The relationship between EFL learners' beliefs and learning strategy use. *System*, *27*, 515-535. http://dx.doi.org/10.1016/S0346-251X(99)00048-2

Yore, L. D., Hand, B. M., & Prain, V. (2002). Scientists as writers. *Science Education*, *86*, 672-692. doi: 10.1002/sce.10042

Zare-ee, A. (2007). The relationship between cognitive and meta-cognitive strategy use and EFL reading achievement. *Journal of Applied Psychology*, *92*(5), 105-119.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 253-269. http://dx.doi.org/10.1037/0033-2909.99.3.432

Biography

Erin Marie Ramirez graduated from Warren Hills Regional High School, Washington, New Jersey, in 2001. She received her Bachelor of Arts in English and Education from Syracuse University in 2005. She was employed as a teacher in Essex County, New Jersey for one year and Middlesex County, New Jersey for two years. She received her Master of Education from Rutgers University in 2011 with a focus in the Social and Philosophical Foundations of Education. Erin was an adjunct professor at George Mason University from 2012-2016 teaching courses in Freshman Orientation, Education, Literacy, and Research Methods.