

**Enriching Teacher Self-Efficacy Through a Support Centric Evaluation Model: A Mixed
Methods Study of TEAM's Impact on Teacher Self-Efficacy**

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Elizabeth Marie Norton
December 2015

Copyright © 2015 by Elizabeth Marie Norton
All rights reserved.

Dedication

I dedicate this dissertation to my mom and dad, Jim and Sue Faller, for always believing that I could do anything. I love you.

Acknowledgements

They say when choosing the people to be in your life, you should choose quality over quantity. So, while this list is small, I am truly grateful to each one of you!

First, a huge thank you to Dr. Angelle, my dissertation chair and trusted advisor, I would have never finished this without you. Your dedication and commitment to helping me through this process has been instrumental to my success.

Matthew, my loving husband, thank you for being my rock and my champion.

Keith, thank you for listening to me talk about this work for hours on end and for sending encouragement when you thought I was losing my mind. You are above proficient at the Internet...and for that I am truly blessed.

Tara and Amy, thank you for helping me find ways to take my mind off the work. Your constant support and encouragement are never wavering. You are the best friends a girl could ask for!

I am looking forward to a future where you all have to call me doc!

Abstract

The effectiveness of teacher evaluation systems is determined by the extent to which they can support improved instructional practice. Research suggests that implementation factors such as attitudes of school leaders (Kimball & Milanowski, 2009), perceptions of fairness (Delvaux, Vanhoof, Tuytens, Vekeman, Devos, & Petegem, 2013), the relationship of the evaluator and the teacher (Weber, 1987), and the quality of the feedback provided can all impact the effectiveness of the evaluation system (Conly & Glasman, 2008; Danielson, 2012; Delvaux et al., 2013; Weber, 1987). This mixed methods study attempted to determine the extent to which these implementation characteristics occurring within the context of the Tennessee Educator Acceleration Model (TEAM) acted as an intervening variable for teacher self-efficacy, the belief system that mediates teacher behavior. Limitations in the sample size resulted in an inability to conduct the statistical analysis needed to determine the extent to which implementation might act as an intervening variable. However, the study did find that the school with the implementation characteristics most aligned to those outlined in the research had an overall higher teacher self-efficacy average on both the pre and post administration of the Teacher Efficacy Belief Scale-Self. The study also examined how the implementation of TEAM influenced teacher efficacy. Bandura (1977) suggests that there are four sources of efficacy development: mastery experience, vicarious experience, verbal persuasion, and affective states. The study found that the teachers' perceptions of fairness and the attitude of the principal gave power to the sources of efficacy. Teachers were more likely to utilize the evaluation experience as a source of efficacy if they perceived the process and feedback to be fair and if there was an expectation that they utilize the process to improve their practice. In addition, teachers were more likely to utilize the feedback provided if it was connected to student outcomes. Support for the evaluation process was also linked to the generation of affective states for the teachers, or to positive and negative stress responses. Finally a model for principal practice is provided that involves the generation of a support centric evaluation model that could serve to ensure that teacher self-efficacy is supported throughout the evaluation process.

Table of Contents

Chapter 1 Introduction to the Study	1
Statement of the Problem	3
Purpose of the Study	7
Research Questions	7
Significance of the Study	7
Definitions	8
Limitations and Delimitations	10
Conclusion	11
Overview of the Study	11
Chapter 2 Literature Review	13
The Search Process	13
The Review Process	14
Theoretical Framework: Self-Efficacy	14
Teacher Self-Efficacy: Construct Evolution Over Time	19
Sources of Teacher Efficacy	30
Effective Evaluation Systems	35
System Characteristics	35
Implementation Characteristics	37
Tennessee Educator Acceleration Model	40
Conclusions	43
Chapter 3 Methodology	45
Research Design	45
Rationale for the Design	46
Role of the Researcher	51
Site and Participants	53
Data Collection Procedures	59
Data Analysis	63
Methods of Verification	70
Conclusion	73
Chapter 4 Analysis and Findings	74
District Context	74
Characteristics of Implementation Context	75
Implementation Summary	92
Research Question 1: To what extent does implementation of TEAM act as an intervening variable for teacher self-efficacy? (QUAN)	92
Research Question 2: How do differences in the implementation of TEAM influence teacher self-efficacy? (QUAL)	95
Conclusion	104
Chapter 5 Discussion and Implications	106
Discussion of Findings in Light of Research on Efficacy	106
A Support Centric Model for Teacher Evaluation	109
Implications for District and State Leaders	116

Recommendations for Evaluation Research.....	116
Recommendations for Efficacy Research	117
Summary.....	118
References	120
Appendix	131
Teacher Efficacy Belief Scale Self-Form.....	132
Teacher Survey Items for Characteristics of Evaluation Implementation.....	134
Teacher Interview Protocol (Original)	135
Teacher Interview Protocol (Revised).....	136
Principal Interview Protocol.....	137
Permission Letter for TEBS-Self.....	138
Implementation Characteristics Survey Results	140
Vita.....	143

List of Tables

Table 1. Research Questions Related to Data Sources.....	50
Table 2. Implementation Characteristics Matrix.....	56
Table 3. Questions Defined by Source of Efficay and by Question Type	65
Table 4. Question Refinement by Data Source	66
Table 5. Question Defined by Evaluation Characteristic and Question Type.....	67
Table 6. Change in Teacher Efficacy for Qualitative Sample.....	68
Table 7. Code Map	72
Table 8. Teacher Self-Efficacy Item Analysis.....	94

List of Figures

Figure 1. Research Design Flowchart	47
Figure 2. Research Design Flowchart Revised.....	48
Figure 3. Support Centric Model.....	112
Figure 4. Support Centric Model Versus Accountability Centric Model.....	114

Chapter 1

Introduction to the Study

A basic Google search of teacher evaluation will result in a range of arguments both for and against teacher evaluation systems across the United States. With federal Race to the Top grants and the Teacher Incentive Fund focused on awarding dollars to those states and districts making substantial changes to their evaluation practices, many states and districts have begun to overhaul evaluation systems to make them both more accurate in measuring teacher performance and more useful as tools for professional growth (Milanowski, 2011). These efforts come on the heels of the New Teacher Project's report, *The Widget Effect*, which documented that 94-99% of teachers in the districts studied scored satisfactory on their performance evaluations (Weisberg, Sexton, Mulhern, & Keeling, 2009). Moreover, data from the increased use of value-added measures has shown varied teacher effects on student achievement (Milanowski, 2011) increasing the need for both more reliable systems of teacher evaluation and systems that can improve teacher practice. The result is two paradigms in teacher evaluation: summative evaluations used for accountability and formative evaluations used for improving teacher performance to increase student outcomes.

The effectiveness of a formative evaluation system hinges on its ability to change teacher practice in a positive way. Research suggests that implementation factors such as evaluator and teacher training (McLaughlin, 1986), attitudes of school leaders (Kimball & Milanowski, 2009), perceptions of fairness (Delvaux, Vanhoof, Tuytens, Vekeman, Devos, & Petegem, 2013), the relationship of the evaluator and the teacher (Weber, 1987), and the quality of the feedback provided can all impact the effectiveness of the evaluation system (Conly & Glasman, 2008; Danielson, 2012; Delvaux et al., 2013; Weber, 1987). Differences in implementation, therefore,

result in differences in how the evaluation system impacts changes to teacher's instructional practice.

The task of changing human behavior is complex. Bandura (1977) explains, "cognitive processes play a prominent role in the acquisition and retention of new behavior patterns" (p. 192). The beliefs that people hold about their abilities, or their self-efficacy, impacts the goals they set, their motivation to reach those goals, and the actions they take (Bandura, 1995). Bandura theorized that one way to change human behavior was by changing the beliefs that mediate that behavior.

Over the past two decades, teacher efficacy research has taken an evolutionary path, evolving from two questions on a RAND study to a complex construct consisting of teacher's self-efficacy beliefs and their outcome expectancies within the context of their teaching tasks (Dellinger, Bobbett, Olivier, & Ellett, 2008). Historically, the two terms, teacher efficacy and teacher self-efficacy have been used interchangeably. Most recently teacher efficacy has been defined as a teacher's belief that his or her actions will have a positive impact on student achievement (Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). The construct defines the teacher's belief that the actions taken in the classroom will overcome external factors resulting in a positive impact on students. Dellinger et al. (2008) deconstructed that construct into two separate belief structures: teacher self-efficacy and teacher's outcome expectancies. Teacher self-efficacy refers to the teacher's belief that he/she can successfully implement best teaching practices (Dellinger, et al., 2008), regardless of whether or not they believe those actions will positively impact students. This separation takes the construct closer to Bandura's original theory of self-efficacy. Bandura (1977) identified these two constructs as outcome expectations and efficacy expectations noting:

Outcome and efficacy expectations are differentiated, because individuals can believe that a particular action will produce certain outcomes, but if they entertain serious doubts about whether they can perform the necessary activities such information does not influence their behavior. (p. 193)

Numerous studies have linked the construct of teacher efficacy to varying indicators of teacher effectiveness. The RAND studies, after the conclusion of a federally funded project, linked high levels of teacher efficacy to the percentage of goals achieved and to the continued use of methods and materials (Armor et al., 1976). Teachers with a strong sense of efficacy have been found to be more receptive to the implementation of new instructional practices (Berman, McLaughlin, Bass, Pauly, & Zellman, 1977; Guskey, 1988), to work longer with struggling students (Gibson & Dembo, 1984), and to persist in the face of challenges and setbacks. Allinder (1994) found that teachers with a strong sense of efficacy spent more time planning and focused more on organization. Research also suggests that a teacher's efficacy has a positive correlation to student achievement (Armor et al., 1976; Ashton & Webb, 1986; Ross, 1992; Cantrell, Almasi, Carter, & Rintamaa, 2012).

However, little research has examined the construct of teacher self-efficacy. If Bandura's (1977) theory about efficacy expectations is correct, then teachers' beliefs in their ability to impact student achievement, or their outcome expectancy, is only part of the belief structure leading to the behaviors they enact. The teachers' beliefs in their abilities to successfully execute those behaviors also play a role in influencing their behavior.

Statement of the Problem

Scant research has been conducted that explores how teacher efficacy is developed, maintained, or modified. Klassen, Tze, Betts, and Gordan (2010) explain:

Although a number of researchers have commented that investigating the sources of teacher efficacy is essential to build a better understanding of how teacher efficacy is formed (e.g., Goddard et al. 2004a, b; Henson 2002), only seven studies –most qualitative or mixed methods- offered empirical explorations of the issue. (p. 31)

Tschannen-Moran, Woolfolk Hoy, and Hoy (1998) suggest that interpretive case studies and qualitative investigations could help “refine our understanding of the process of developing efficacy” (p. 242).

Despite the lack of clarity in how efficacy is developed and improved, many states have begun to instate teacher evaluation systems aimed at improving teacher practice. Bandura (1977) suggests that the behaviors people enact are mediated by their self-efficacy beliefs. So, in order to change teacher behaviors, evaluation systems may need to first change teacher self-efficacy. However, evaluation systems created to generate changes to teacher self-efficacy and result in changes in teacher behavior cannot support that contention without a more thorough understanding of how self-efficacy is developed.

One example comes from the state of Tennessee, which has recently overhauled its state evaluation system, implementing the Tennessee Educator Acceleration Model (TEAM) for teacher evaluation. The evaluation model is intended to serve as both a summative accountability measure and a formative tool to improve teacher effectiveness. The Tennessee Department of Education states on the TEAM website (team-tn.org):

The Tennessee Educator Acceleration Model (TEAM) is about principals and teachers working together to ensure the best possible instruction everyday. Through frequent observation, constructive feedback, student data, and professional development, TEAM is

designed to support all educators in doing their best work to help every student learn and grow. (para. 1)

However, little research exists examining how the evaluation system works to create changes in teacher practice. Connor (2013) states:

More work needs to be accomplished [in teacher evaluation] and these observation tools need continuous investigation and refinement to support effective teaching- we still do not understand everything that it takes to improve students' outcomes. However these [evaluation] tools represent an important step forward and a move toward policy that promises to make a true difference in what is defined as high quality and effective teaching, what it looks like in the classroom, and how these practices can be widely disseminated so that all children, including those attending under resourced schools, can experience effective instruction, academic success, and the lifelong accomplishment that follows. (p. 345)

Currently many teachers and principals are requesting changes be made to the evaluation system in Tennessee. An article on the Michigan Association of School Administrators website (<http://gomasa.org/news/tennessee-struggles-get-teacher-evaluation-right>), states that while Tennessee education officials defend the new system, principals have concerns. Grover Whitehurst of the Brookings Institution is also quoted, saying, "There's a lot we don't know about how to evaluate teachers reliably and how to use that information to improve instruction and learning" (para. 6).

In fact in one district in Tennessee alone, a recent survey of teachers demonstrated varying teacher perceptions about the influence of the system on their teaching. Some of these comments included:

- I feel that TEAM evaluations are too stressful and too overwhelming for teachers to be effective teachers.
- Evaluations are very useful in providing a teacher with feedback about his/her instruction, improvements necessary to make instruction more effective, and encouragement to continue using strategies that are best practices.
- I'm working harder than ever and the TEAM rubric is very unfair.
- I think the TEAM rubric is a good tool to use to guide instruction.
- This eval system has lowered morale and made everyone feel ineffective.
- All of my evaluations have helped me become a better teacher. I have always got great feedback that has helped improve instruction.
- The Team model of evaluation is totally unfair to teachers and needs to be replaced (accessed from knoxschools.org on January 25, 2014).

These perceptions are an indicator that the implementation of the system varies from school to school, even within a single district. However, it is these perceptions that are driving changes in the evaluation system at the state level. Education Partnerships, Inc. (2011) defined effective evaluation systems as ones that promote a positive school climate, increase the principal's instructional leadership, and stimulate professional development. Danielson (2012) also stated that evaluation procedures should promote professional development, while also ensuring valid and reliable results. However, some evaluation systems, such as TEAM, are designed to be a catalyst for improved teacher performance in and of themselves. Aimed at improving teacher behaviors, the systems are meant to also improve teacher self-efficacy, the beliefs that mediate teacher behavior. Redefining effectiveness of the system and linking it to a change in teacher

self-efficacy, could aid in understanding just how effective these systems are at reaching their intended outcomes.

Purpose of the Study

In an effort to understand how implementation of TEAM, an evaluation model designed to support effectiveness in teacher practice, impacts teacher self-efficacy, the belief system mediating behavioral change, a mixed method study will be conducted. The purpose of the study is to explore the extent to which the implementation of the evaluation system acts as an intervening variable for teacher self-efficacy. In addition, the study will examine how the implementation of TEAM influences teacher self-efficacy by examining the characteristics of implementation through the lens of Bandura's four sources of efficacy: mastery experience, vicarious experience, verbal persuasion, and affective states.

Research Questions

In this examination of whether TEAM impacts teacher self-efficacy the following research questions will be explored:

1. To what extent does implementation of TEAM act as an intervening variable for teacher self-efficacy? (QUAN)
2. How do differences in the implementation of TEAM influence teacher self-efficacy? (QUAL)

Significance of the Study

By examining whether Tennessee's evaluation system impacts teacher self-efficacy, a picture of how efficacy is developed and supported by an evaluation model might be provided. Clarity for how these beliefs are changed might also be provided, contributing to the research on

efficacy development. Current research on efficacy development is limited by a lack of quantitative data to demonstrate the extent of the changes observed, or by loosely defined teaching tasks tied to changes in efficacy beliefs. By exploring efficacy through clearly defined teaching tasks related to best practices in instruction this research might provide researchers with a more concrete lens for defining teacher self-efficacy and exploring its development in the future.

This picture of efficacy development might help Tennessee policy makers as they continue to refine the state evaluation model. Understanding how differences in implementation characteristics impact efficacy development could assist in knowing where to leverage resources or provide more professional development to district and school leaders to make the evaluation system work. In addition, providing another way to define effectiveness for an evaluation system, might aid in clarifying how the system is working towards the goals it was intended to reach. Knowing the extent to which the system is impacting teacher self-efficacy could provide policy makers with additional evidence for making future decisions about the system.

Understanding how differences in implementation characteristics impact teacher self-efficacy could also be beneficial for district and school leaders as they seek to implement the state evaluation model. Knowing which characteristics of implementation impact teacher self-efficacy either positively or negatively could provide a map for district and school leaders to improve upon their use of the system and their own practices related to the evaluation process.

Definitions

1. Affective states: Levels of emotional arousal in high stress situations. Emotions such as excitement or anxiety add to an individual's or a group's perceptions of their ability to cope in certain situations (Goddard, Hoy, & Woolfolk Hoy, 2004).

2. Mastery experience: An experience where one is successful in completing a task.

Research findings have suggested that mastery experiences raise efficacy levels whereas failure at a task tends to lower efficacy (Bandura, 1977).

3. Self-efficacy: An individual's belief in their abilities. Bandura (1977) theorized that a person's belief in his ability to successfully complete a task is directly related to the behaviors he will exhibit in completing the task.

4. Teacher efficacy: A teacher's beliefs in his or her abilities to impact student outcomes. These beliefs influence behaviors such as planning, effort in the classroom, and coping strategies utilized when students struggle (Goddard, Hoy, & Woolfolk Hoy, 2004).

5. *Teacher self-efficacy*: A teacher's beliefs in his or her ability to successfully perform specific teaching tasks in a given context (Dellinger, Bobbett, Olivier, & Ellett, 2008).

6. Verbal persuasion: Verbal encouragement that one can be successful in completing a task. Social persuasion can strengthen an individual's or faculty's conviction that they are capable of completing a task or meeting a goal (Goddard, Hoy, & Woolfolk Hoy, 2004). Although efficacy beliefs increased in this manner tend to be weaker than those raised through mastery experiences, they can contribute to an increased effort in completing a task. This effort could in turn produce a mastery experience that will strengthen efficacy beliefs even more (Bandura, 1977).

7. Vicarious experience: Seeing or hearing about a similar person's success in completing a task. Seeing others successfully perform a task can generate expectations that the observer is also capable of completing the task. Vicarious experience relies on inferences being made from social experiences in order to raise efficacy (Bandura, 1977).

Limitations and Delimitations

Delimitations of the Study

This study was restricted to elementary teachers in one district in the southeastern United States, thus, limiting the generalizability of the findings to elementary teachers. In addition, the study was delimited to a rural school district. Findings may not be generalizable to urban or suburban districts. The study was delimited to schools utilizing the rubric created by the National Institute for Excellence in Teaching (NIET) and whose evaluators had been certified by NIET, which in this study included one school utilizing The System for Teacher and Student Advancement (TAP) evaluation process and four utilizing TEAM. The system characteristic of being both summative and formative, the structure for meaningful feedback and collaboration, and the multiple measures of teacher performance are identical in both systems. The detailed criterion for measuring teacher performance has a slight variance in wording in several indicators: standards and objectives, presenting instructional content, lesson structure and pacing, questioning, academic feedback, and thinking. These variances were added to the TEAM rubric in order to clarify the language. Including the TAP school in the sample increased the variability in the implementation characteristics. However, with only one TAP school represented, these findings are not generalizable to TAP schools or to any schools utilizing a different evaluation system.

Limitations of the Study

The survey instrument used required self-reporting which limits the validity of the results. Self-report instruments are dependent upon the individual giving an honest representation of their beliefs. However, sometimes respondents respond in the way they believe is either most socially acceptable, or in the way they believe the researcher wants them to respond. This could

cause the data to be invalid (Gay, Mills, and Airasian, 2009). In addition, the researcher is involved in coaching and supporting schools in their use of the TAP system, which utilizes an evaluation process with the same system characteristics as TEAM. The researcher believes that the evaluation model can be a useful tool for improving teacher effectiveness when implemented well. In order to account for this bias, the researcher will provide a transparent report of the data analysis demonstrating how codes and themes were generated from the qualitative data. The researcher will also maintain a reflexive journal to document decisions made within the data collection and analysis processes.

Another limitation is time. Data collection occurred in one semester. This limits the amount of potential change to only that which may be observed in a short time. Behavioral changes may take longer to generate than what might be captured in one evaluation cycle. Examining the changes over longer periods of time may be necessary to fully understand the potential impact of the evaluation process on teacher beliefs.

Conclusion

Evaluation can provide a catalyst for improvements in teacher practice. Understanding how one evaluation system influences teacher self-efficacy might add clarity to our understanding of how efficacy is developed, as well as to how evaluation systems work to encourage changes in teacher practice. As this study explores how the Tennessee Educator Acceleration Model influences teacher self-efficacy, it may add to clarity to the definition of teacher self-efficacy as a construct.

Overview of the Study

This mixed method study will examine the extent to which TEAM acts as an intervening variable for teacher self-efficacy. It will explore how implementation of TEAM influences

teacher self-efficacy utilizing Bandura's four sources of efficacy as a framework for examination.

Chapter 2 will contain a literature review of theoretical framework of self-efficacy followed by literature surrounding the evolution of the construct of teacher efficacy. Then a review of the literature on how efficacy is developed will be offered. This will be followed by a review of the literature defining effective evaluation systems for teachers. Finally, a review of the literature for the TEAM model for evaluation will be provided. Chapter 3 will include an explanation of the methodology and procedures used for this mixed methods research study. In Chapter 4, a presentation of the findings will be provided. The study will conclude in Chapter 5 with a discussion of the findings and their implications for future research.

Chapter 2

Literature Review

This research addresses the extent to which the TEAM system impacts teacher self-efficacy and seeks to answer the following questions:

1. To what extent does the implementation of TEAM act as an intervening variable for teacher self-efficacy? (QUAN)
2. How does the implementation of TEAM influence teacher self-efficacy? (QUAL)

This chapter begins with a synopsis of the search process used in this review. The literature review then provides an overview of Bandura's theory of self-efficacy. This is followed by a review of the current research in education centered on teacher self-efficacy and the development of the construct over time. Next a review of the literature on the implementation characteristics of effective evaluation systems is discussed. Finally a review of the empirical literature surrounding the TEAM system is provided.

The Search Process

Sources for this literature review were found by searching multiple databases. The initial search for books and articles related to teacher efficacy and self-efficacy was completed at the University of Tennessee Library as well as its interlibrary loan resource. Additional searches using ERIC, Education Full Text, and Google Scholar were then conducted using the search terms self-efficacy, teacher efficacy, teacher self-efficacy, and efficacy. These databases were utilized due to their accessibility and timeliness (Anderson & Kanuka, 2003). A search for books on these topics was conducted on Amazon. Additional sources were found by locating sources listed by the authors of the articles and books found through the initial search methods. Additional resources were also acquired through professors and colleagues.

The same search process was utilized for the review of literature on teacher evaluation and the TEAM system. The search terms teacher evaluation, evaluation, TEAM, Tennessee Educator Acceleration Model, and teacher observation were used for the initial search.

The Review Process

The teacher efficacy literature was uploaded into Mendeley for reading and annotating. This allowed articles to be organized by topic and searches to be conducted on key terms or notes. The literature was then placed within the historical context in which it was conducted as well as compared to other similar studies (Boote & Beile, 2005). Following the initial annotation process, the articles on teacher efficacy were then uploaded into ATLAS.ti where they were coded for organizational themes. Rereading the literature within the coded themes allowed for greater synthesis of the literature and a deeper understanding of the literature as a whole. This also aided in clarifying what had already been done within efficacy research and distinguishing it from what research still needed to be conducted (Boote & Beile, 2005).

For the review of literature on teacher evaluation and TEAM, articles were loaded into a folder on the computer desktop. Articles were read and annotated in preview. Organizational categories were generated and recorded on a white board. Articles that provided evidence for each category were then listed beneath the organizational heading.

Theoretical Framework: Self-Efficacy

One way that human action is created and maintained is through the relationship between knowledge and action. This relationship is mediated by a person's self-efficacy, or their belief in their own ability to complete a task. A variance in a person's self-efficacy can cause a person to either give up when faced with an obstacle or to persist when faced with the same obstacle (Bandura, 1982). In order to test this theory, Bandura (1982) employed a microanalytical

methodology where individuals were “presented with graduated self-efficacy scales representing tasks varying in difficulty, complexity, stressfulness, or some other dimension” (p. 123).

Bandura then analyzed the degree to which a person’s self-efficacy was congruent with the action level of a given task. Bandura examined efficacy induction by having his subjects perform tasks gaining either mastery experiences themselves or observing others gain mastery experiences and then indicating their perceived self-efficacy before completing the task. He found that a person’s perceived self-efficacy was a better predictor of future performance than performance attainment.

Other studies have similar findings about the relationship between self-efficacy and performance attainment. For example, in a study of 29 male and 34 female participants who had recently quit smoking, DiClemente (1981), found that participants who were able to maintain smoking cessation at the 5-month follow-up had significantly higher initial self-efficacy scores ($p < .005$) than recidivists, or those participants who returned to smoking. Bandura and Adams (1977) also found a correlation between performance attainment and self-efficacy ($p < .01$) in a study of 10 subjects with chronic snake phobias. In the study, subjects were provided a systematic desensitization treatment for their phobia. The subjects were tested prior to the treatment and following the treatment to determine their levels of self-efficacy, as well as their level of approach behavior towards the snakes. At the end of the study, those with higher levels of self-efficacy had higher levels of approach behavior. Only one of the subjects was able to complete the tasks with desensitization alone. Eight of the other nine subjects were provided modeling from the therapist. Following the model, all 8 were able to successfully complete the performance tasks. Bandura and Adams also calculated a microanalytic measure of congruence

by computing the percent of equivalence between the subjects' self-efficacy and task performance on each individual task (84% congruence).

Understanding that self-efficacy is a strong predictor of performance, Bandura (1977) postulated four sources of efficacy that could lead to behavioral changes: mastery experience, vicarious experience, verbal persuasion, and affective states. The sources are not isolated, as "any given method, depending on how it is applied, may of course draw to a lesser extent on one or more other sources of efficacy information (p.195)." Rather, they work together to build a person's belief system. Once established, change in self-efficacy transfers to other situations where a person once held lower efficacy beliefs. So together, the four sources of efficacy form a conceptual framework for the study of behavioral change.

Mastery Experience

Mastery experience is the source of efficacy that is the most influential factor in efficacy development (Goddard, Hoy, & Woolfolk Hoy, 2004). When a person successfully completes a task, efficacy is raised. However, when a person believes they have been unsuccessful, efficacy tends to lower (Bandura, 1977). Both the timing of the experience and the total pattern of success and failure determine the effects of a mastery experience. Early failures may have a stronger negative influence on efficacy than an occasional failure that is set in the midst of a string of successes.

The field of psychology utilizes several treatments that Bandura (1977) categorizes as creating mastery experiences. One of these is the participant modeling approach. In this treatment, clients are exposed to the feared behavior in a controlled environment where they are exposed at graduated intervals to the task and may perform the task jointly with the therapist. The aids are gradually removed until the client has a mastery experience that is unassisted.

Vicarious Experience

Seeing others succeed at difficult tasks can cause observers to believe that they too can succeed if they put forth enough effort (Bandura, 1977). Vicarious experiences are achieved when one sees another person complete a task. Just as with mastery experience, seeing someone fail at a task can have negative effects on efficacy and the timing of the experiences, as well as its placement in a pattern of success or failure plays a part in the experience's impact on efficacy. In addition, the model observed has some bearing on the potential impact of a vicarious experience. The more similar the model is to the observer, the more impactful the model is to a person's efficacy. So, a model with similar characteristics that performs a task successfully is more likely to impact an observer's beliefs in their capabilities than a model that is dissimilar. In addition, behaviors that have clear outcomes convey more efficacy information than modeled behaviors with ambiguous outcomes. Vicarious experiences are also more impactful when they are demonstrated by a variety of models.

As previously mentioned, Bandura and Adams (1977) found a correlation between increases in self-efficacy and approach behaviors in people with snake phobias. In this same research, Bandura and Adams were also able to show how mastery and vicarious experiences can affect participants' efficacy and subsequently their behaviors. Some of the participants were guided through mastery experiences with snakes, while others merely observed the therapist perform similar activities with the snakes. They measured the efficacy expectations of the participants before and after each treatment and found that mastery experiences produce higher efficacy expectations than vicarious experiences. However, both experiences produced gains in efficacy.

Verbal Persuasion

Verbal persuasion involves leading people through suggestion into believing that they are capable of completing tasks successfully. Efficacy beliefs that are created through verbal persuasion are weaker than those that are built through mastery experience. However, verbal persuasion can increase the impact of a mastery experience when the two are combined together (Bandura, 1977).

In a study of male and female students at North Texas State University, Weinberg, Gould, and Jackson (1979) found that the use of verbal persuasion had a significant impact on the students' performance on a muscular endurance task ($p < .001$). Prior to the experiment there were no differences in estimates of how long students could extend their leg. Students were then assigned randomly to either a group that would receive feedback that the opponent had weak abilities or to a group that would receive feedback that the opponent had strong abilities. Students assigned to the weak opponent performed better than students assigned to the strong opponent and reported higher expectations for their performance prior to competing. In addition, the mastery experiences of each group affected their second performance of the task. Students in the high efficacy group extended their legs for longer periods of time on the second trial, whereas low efficacy students extended their legs for shorter periods of time on the second trial.

Affective States

According to Bandura (1977), "Emotional arousal is another constituent source of information that can affect perceived self-efficacy in coping with threatening situations" (p. 198). The way that a person copes with and handles stress has a significant impact on their motivation to complete a task, as well as their belief that they can successfully complete that task. Performance expectations can be negatively impacted when stress causes a negative emotional

response. However, eliminating negative responses may not be enough to cause a change in the self-efficacy of an individual (Bandura & Adams, 1977).

Teacher Self-Efficacy: Construct Evolution Over Time

In examining the research on teacher efficacy, it is important to note that two theories for behavioral change have been instrumental in informing the direction of teacher efficacy research. As a result, the construct of teacher efficacy has been defined and redefined continuously since its original inception.

While Bandura's theory of self-efficacy will be the primary basis for this research, Rotter's Theory of Behavioral Change was the catalyst for teacher efficacy research in the mid-1970s (Labone, 2004; Hoy & Spero, 2005; Tschannen-Moran, Hoy & Woolfolk Hoy, 1998; Wheatley, 2002;). Rotter's theory examines internal versus external locus of control as a catalyst for human behavior (Rotter, 1966). Rotter suggests that a person's beliefs about the nature of the causal relationship between one's own behavior and the consequences of those behaviors affect a variety of future behavioral choices. For teachers, the belief that their own behaviors have an impact on student achievement, is characterized as an internal locus of control, whereas the belief that outside forces, such as home life or socio-economic status are more influential is characterized as an external locus of control.

In 1975, the RAND Corporation was contracted to conduct a study "whose purpose [was] to identify the school and classroom policies and other factors that have been most successful in raising the reading scores of inner-city children" (Armor et al, 1976, p. 5). The RAND Corporation selected a sample of 20 elementary schools in the Los Angeles Unified School District for the study. Using Rotter's theory as a base, the group included a survey that included two items relating to what they termed 'efficacy':

1. When it comes right down to it, a teacher really can't do much (because) most of a student's motivation and performance depends on his or her home environment.
2. If I try really hard, I can get through to even the most difficult or unmotivated students. (p. 33)

The result was a finding that marks the beginning of teacher efficacy research: Teachers who felt 'efficacious' were associated with observed gains in reading performance.

The RAND Corporation published a second study in May of 1978. This study examined what happens to local projects in ESEA Title III and ESEA Title VII programs when federal funding ends. Participants from 18 school districts and a total of 100 sites were selected to participate. Utilizing the same two questions about teacher efficacy, the second RAND study also found that teacher sense of efficacy showed strong positive effects on all outcomes.

Following the RAND studies, teacher efficacy research began to diverge into two different theoretical backgrounds. One group leaned heavily on Rotter's Social Learning Theory, while the other took on the theoretical perspective of Bandura's Social Learning Theory. The result was a variety of definitions and measures for teacher efficacy and an ever-evolving construct. The following review of literature will examine the evolution of the construct that led to the final definition by Dellinger et al. (2008) that teacher self-efficacy is a teacher's beliefs in his or her ability to successfully perform specific teaching tasks in a given context.

Initiated by the RAND studies' findings that reported a significant relationship between teachers' sense of efficacy and student achievement, utilizing a grounded theory approach, Ashton, Webb, and Doda (1983) worked to "develop a conceptual framework for understanding the nature, antecedents, and consequences of efficacy attitudes in teachers and to suggest further

research necessary to reject, elaborate, and/or extend the conceptual framework” (p. 9). They examined how teaching efficacy, the teacher’s beliefs about the relationship between teaching and learning, and personal efficacy, the teacher’s general sense of effectiveness as a teacher, worked to create personal teaching efficacy. This preliminary conceptual framework was used to design the first phase of the study where 49 teachers from two schools were asked to respond to a questionnaire about their feelings about teaching. Four teachers were then selected for observations and interviews. During the second phase of the study, Ashton, Webb, and Doda applied findings from phase one to interviews of 48 high school basic skills teachers. During this phase, a process-product study and a pilot study comparison of three approaches to increase teacher efficacy were also conducted. While Ashton, Webb, and Doda found a significant positive correlation between teachers’ sense of efficacy and student achievement, they cautioned against viewing teacher efficacy as a stable variable:

Our research suggests that teachers’ sense of efficacy is reciprocally and multiply determined by a complex and interrelated system of variables. An adequate understanding of the dynamics affecting teachers’ sense of efficacy requires a perspective that reflects the complexity of the relationships existing among these variables. Traditional research approaches that assume a linear, additive model and conceive effects in terms of antecedents and consequences are inadequate for the task of discovering the complexity of the relationships existing in regard to teachers’ sense of efficacy. (p. 27)

In concluding their study, Ashton, Webb, and Doda stated that “teachers’ sense of efficacy is negotiated daily in their myriad transactions with students, peers, and administrators” (p.38). They suggested that rather than focusing on the identification of teacher efficacy as a character

trait, future research should examine the “socialization practices, organizational structures, instructional techniques, administrative strategies, and home-school relations” (p.37) that might work to support teachers’ sense of efficacy. Ashton, Webb, and Doda defined the construct as the extent to which teachers believe they can impact student learning.

To examine teacher efficacy against various variables and build on the work of Ashton, Webb, and Doda, Gibson and Dembo (1984) expanded the two items from the RAND studies to create a 30-item measure of teacher efficacy that employed Bandura’s Social Learning Theory. The instrument was created beginning with a pilot study where 53 sample items were administered to 90 teachers, items were eliminated that either had poor variability or did not load clearly on one of the substantial factors for teacher efficacy. Gibson and Dembo then utilized the measure in a three phase study of teacher efficacy. In the first phase, 208 elementary school teachers from 13 elementary schools completed the 30-item Teacher Efficacy Scale and a factor analysis was completed on the data. Three questions were the focus of this phase: “What are the dimensions of Teacher Efficacy? How do these dimensions relate to Bandura’s theory of self-efficacy? What is the internal consistency of the teacher efficacy measure?” (Gibson & Dembo, 1984, p. 573). According to Gibson and Dembo, “two substantial factors emerged from the factor analysis, with Factor 1 accounting for 18.2% of the total variance and Factor 2 accounting for 10.6% of the total variance” (p. 573). Factor 1 became known as personal teaching efficacy, or PTE and factor two general teaching efficacy, or GTE (Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). Gibson and Dembo stated that the first factor, personal teaching efficacy relates to a teacher’s belief that he/she has the skills necessary to impact student achievement and they related this to Bandura’s definition of self-efficacy. Factor 2, general teaching efficacy, is said to relate to a teacher’s belief in the general ability of a teacher to influence achievement versus the

impact of outside factors such as home environment or family background, and Gibson and Dembo link this to Bandura's definition of outcome expectancy. Tschannen-Moran, Woolfolk Hoy, and Hoy (1998) later made a case for why this factor is different than outcome expectancy:

Bandura (1986) argued that an outcome expectancy is a judgment of the likely consequences of a specific action, given an individual's anticipated level of performance (a means-end relationship as described by Skinner, 1996). Bandura pointed out that outcome expectancy adds little to the explanation of motivation because the outcome a person expects stems from that person's assessment of his or her own capabilities and expected level of performance, not from what it would be possible for others to accomplish under similar circumstances. Therefore, the items used to measure the second factor of teacher efficacy, about the possible impact of teachers in general (GTE), cannot be considered an outcome expectancy. (p. 223)

In phase 2 of their study, Gibson and Dembo asked the following questions: "Does evidence of teacher efficacy gathered from different sources in different ways converge? Can teacher efficacy be differentiated from other constructs?" (p. 574). To answer those questions, they selected fifty-five teachers enrolled in graduate courses at a state university in California. The teachers completed two measures, one close ended and one open ended, for each of the following constructs associated with effective teachers: teacher efficacy, verbal ability, and flexibility. The results verified the difference between teacher efficacy and the two other constructs, as well as the convergence of teacher efficacy when measured by the two different approaches, suggesting that the Teacher Efficacy Scale could be used to measure teacher efficacy.

Phase 3 consisted of classroom observations of 8 teachers selected from the 208 teachers from phase 1 of the study, 4 with high efficacy and 4 with low efficacy. In this phase 3 the following question was investigated: “Do high and low efficacy teachers exhibit differential patterns of teacher behaviors in the classroom related to academic focus, feedback, and persistence in failure situations?” (p. 576). Gibson and Dembo found that the high efficacy teachers spent less time in small group instruction and more time in whole group. They also found that the high efficacy teachers spent more time monitoring and checking seat work than the low-efficacy teachers. The high efficacy teachers also spent more time in preparation or paperwork than the low-efficacy teachers. Low-efficacy teachers were observed giving criticism when students responded incorrectly to questions. However, no criticism was observed in the classrooms of high-efficacy teachers. High-efficacy teachers also provided more wait time to students when questioning and led students through questioning to find correct answers, whereas low-efficacy teachers were more likely to provide the answer and move on. While the sample size for this phase of the study is limited, it does suggest that differences in efficacy do result in differences in teacher behavior in the classroom.

Gibson and Dembo’s (1984) work confirmed the value of studying teacher efficacy as a construct. It also added multiple dimensions to the definition of teacher efficacy. Teacher efficacy was now defined as an “integration of teaching efficacy and personal teaching efficacy” (Dembo & Gibson, 1985). Dembo and Gibson (1985) elaborate on the complexity of the model saying, “There is some question, however, concerning whether the general component of teacher efficacy is essential to a model of teacher efficacy” (p. 175).

In an effort to further clarify the construct, Woolfolk and Hoy (1990) utilized the Gibson and Dembo instrument in a study of 182 liberal arts majors enrolled in a teacher preparation

program examining the efficacy of pre-service teachers. The study included 20 of the original 30 items from the Gibson and Dembo instrument: the 16 that produced adequate reliability and 4 that referred to the adequacy of the teacher's pre-service preparation. Woolfolk and Hoy also included the 2 original RAND items for a total of 22 questions on the survey administered to pre-service teachers. The study sought to answer two questions: "Is the structure of efficacy for prospective teachers the same as has been found for experienced teachers?" and "Are prospective teachers' beliefs about efficacy related to their orientations toward discipline, order, control, and motivation in schools?" (p. 81).

Upon analysis of the efficacy measure, Woolfolk and Hoy found that the same two factors: Personal Teaching Efficacy and General Teaching Efficacy emerged, as they had in the original Gibson and Dembo (1984) study. In addition, the findings of the study suggested that Personal Teaching Efficacy could also be divided into two related aspects of responsibility for positive student outcomes and responsibility for negative student outcomes.

Following the Woolfolk and Hoy study, Guskey and Passaro (1994) also utilized the Gibson and Dembo instrument to "bring clarity to interpretations of teacher efficacy measures" (p. 630). Guskey and Passaro noticed that the items that loaded on personal efficacy all used the referent I, and that all the items loading on teacher efficacy all used the reference "teachers." In addition they noticed that all the personal efficacy factors had positive presuppositions (i.e., "I can.") and all the teacher efficacy factors had negative presuppositions (i.e. "teachers cannot."). This led them to question whether the structure of efficacy was truly based on the two factors of personal and teaching efficacy, or if it was simply internal versus external locus. Hoping to "broaden our understanding of this important construct and improve our means of measuring it by unraveling the factors that determine its structure" (p. 632) Guskey and Passaro administered

an altered form of the Gibson and Dembo instrument to 283 experienced classroom teachers and 59 preservice teachers. Guskey and Passaro describe the construction of the instrument:

We began with 16 items that yielded significant factor loadings in Gibson and Dembo's (1984) original study, 15 of which were employed in Woolfolk and Hoy's (1990) extended study. To these we added the three additional items that Woolfolk and Hoy had found to yield significant factor loadings, plus the two Rand items. Of these 21 items, 12 had been found to load principally on the personal efficacy dimension and 9 on the teaching efficacy dimension...Items were then altered in the following way: Seven of the 12 personal efficacy items, all of which reflected personal-internal orientation (P-I) were randomly selected and reworded to reflect either a teaching-internal (T-I) or personal-external (P-E) orientation...In a like manner, four of the nine teaching efficacy items, most of which reflected a teaching-external (T-E) orientation were randomly selected and reworded to reflect either a personal-external (P-E) or teaching-internal (T-I) orientation. (p. 633)

The result of the study was an even more complex definition of teacher efficacy. Teacher efficacy was defined as being more connected to beliefs about internal and external locus of control. Guskey and Passaro (1994) stated, "today's most widely accepted scales of teacher efficacy may be measuring beliefs about the relative contribution of internal and external factors to student learning and performance" (p. 640) and they suggested that further studies refine the definition of the construct.

In 1998, Tschannan-Moran, Woolfolk Hoy, and Hoy examined "the theoretical and empirical underpinnings of teacher efficacy" (p. 202) by examining the results of various studies that utilized the Gibson and Dembo instrument. The study sought to clarify the construct of

teacher efficacy and improve its measurement. The result of their study was the introduction of a model of teacher efficacy that merged the two competing conceptual strands in the literature: the strand based on Rotter's work and the one based on Bandura's work. Tschannan-Moran, Woolfolk Hoy, and Hoy suggested that teacher efficacy is dependent upon circumstances or environment, that the setting, students, and subjects being taught all play a role in the efficaciousness of a teacher. They defined teacher efficacy as a teacher's belief in his or her ability to execute a specific teaching task successfully within a given context. This definition is distinctly different from previous definitions of teacher efficacy because it focuses in on the teacher's analysis of a given teaching task in light of the context in which it will be performed, rather than a focus on the teacher's beliefs about being able to overcome external factors to impact student outcomes. The connection between the teaching task and anticipated outcomes for students was still prevalent in this model. Tschannen-Moran, Woolfolk Hoy, and Hoy described these two elements of teacher efficacy as analysis of the teaching task, where student ability and motivation are a factor, and assessment of personal teaching competence.

Dellinger, Bobbett, Olivier, and Ellett (2008) expanded on the idea of clarifying the teaching task and further clarified the differences between teacher efficacy and teacher self-efficacy. They defined teacher self-efficacy as "teachers' individual beliefs about their own abilities to successfully perform specific teaching and learning related tasks within the context of their own classrooms" (p. 751). They made a clear distinction between teacher efficacy and teacher self-efficacy as two different constructs, defining teacher efficacy as a "teachers' beliefs in their abilities to affect student performance" (p. 753). In an effort to provide a clearer measure of teacher self-efficacy, Dellinger, Bobbet, Olivier, and Ellett created the Teacher Efficacy Belief Scale-Self or TEBS-Self. This new measure reflected the meaning of teacher self-efficacy as a

belief in one's own abilities to successfully complete specific teaching tasks. The measure used tasks that are meaningful in terms of effective teaching and learning. Examples include rating beliefs in ability to provide appropriate feedback to students, ask higher order questions, and communicate learning outcomes to students. The measure drew on the language from the PACES classroom observation framework for the assessment of teaching and learning. The PACES framework has similar language to other teacher evaluation instruments like the TAP Rubric (National Institute for Teaching and Learning, 2013) and The Framework for Teaching Evaluation Instrument (The Danielson Group, 2013), all of which use language that describes research-based best practices in teaching.

The measure was tested for validity and reliability in three phases. In the first phase, Dellinger et al. (2008) used three different forms that utilized three different item stems. The team gave the three measures to 434 teacher volunteers with each teacher responding to two of the three forms. The BELIEF (*my belief in my ability to...is...*) elicited different results than the more traditional stems and was consistent with the language of self-efficacy. As a result, that stem was adopted for all subsequent versions of the TEBS-Self.

In phase two, 51 items were developed utilizing the PACES classroom observation framework for the assessment of teaching and learning as the selection criteria for meaningful teaching tasks. Phase three involved a panel of 45 professional educators who ranked the 51 items according to importance of each task in assessing teacher self-efficacy. The 30 items with the highest ratings were chosen for the final item pool. The items were then put with a four point Likert scale: "1 = very weak belief in my capabilities, 2 = moderate belief in my capabilities, 3 = strong belief in my capabilities, and 4 = very strong belief in my capabilities" (Dellinger et al., 2008, p. 757). The measure was then used in 3 separate dissertation studies.

Teacher self-efficacy as defined and measured by Dellinger et al. (2008) is a construct that more closely aligns to Bandura's theory of self-efficacy by aligning the beliefs with specific teaching tasks in the current teaching situation. This is different from teacher efficacy, which has historically been more connected to a belief that the teacher can impact changes in student achievement. Teacher self-efficacy is looking at just the teacher's belief that he/she can successfully complete the task and those tasks are defined as the tasks necessary to be an effective teacher.

In examining the research on teacher efficacy and teacher self-efficacy, it is important to note this distinction in definition. There is abundant research linking teacher efficacy to student achievement (Armor et al., 1976; Ashton & Webb, 1986; Cantrell, Almasi, Carter, & Rintamaa, 2012; Moore & Esselman, 1992; Ross, 1992). Within these studies, the term teacher efficacy and teacher self-efficacy are sometimes used synonymously. However, these studies are examining the teacher's beliefs in their ability to impact student achievement.

Some studies have linked teacher self-efficacy and student achievement using Bandura's Teacher Self-Efficacy Questionnaire. For example, in a study of 1,043 fifth grade students, Guo, Conner, Yang, Roehrig, and Morrison (2012) found that teacher self-efficacy, as measured by Bandura's (1997) Teacher Self-Efficacy Questionnaire, was positively correlated to fifth grade overall literacy scores ($p = .02$), noting that, "students whose teachers had higher self-efficacy were more likely to obtain higher scores in literacy" (p. 16). Bandura's instrument measures efficacy to influence decision-making, efficacy to influence school resources, instructional self-efficacy, disciplinary self-efficacy, efficacy to enlist parental involvement and community involvement, and efficacy to create a positive school climate. While it does contain an instructional self-efficacy component, the items within that section are very similar to items

found on other measures of teacher efficacy. For example, “How much can you do to promote learning when there is a lack of support at home?” is an item on the Bandura instrument that is highly similar to items found on other efficacy measures (p. 1). In addition, the items do not capture the teacher’s belief about their ability to effectively implement best teaching practices.

However, no studies were found that utilized the TEBS-Self to examine the correlation between teacher self-efficacy and student achievement. A search for the TEBS-Self in Eric via ProQuest led to only the Dellinger et al. (2008) introduction of the measure. A subsequent search on Google led to several international studies that utilized the instrument either as the grounds for creation of a new instrument (Ceylandag, 2009; DeChenne, Enochs, & Needham, 2012) or to measure self-efficacy against other variables (Holt, 2011; Karimvand, 2011; Celik, 2013).

Sources of Teacher Efficacy

Bandura (1977) theorized that self-efficacy was created and maintained through four sources: mastery experiences, vicarious experiences, verbal persuasion, and affective states. According to Klassen et al. (2010), few studies have investigated the sources of teacher efficacy and how teacher efficacy is developed.

Mastery Experiences

Mastery experiences in education can be defined as classroom experiences where teachers have experienced success. Many studies have analyzed mastery experiences through the lens of how the teacher defines success. For example, Milner (2002) conducted a case study on one European American high school teacher. In the study, the teacher defines mastery as getting tougher in her teaching and grading, and measured it a success based on the verbal feedback provided to her by students, parents, and colleagues. Milner and Woolfolk Hoy (2003) labeled the teacher’s desire to “demystify or change negative stereotypes about African Americans” (p.

268) as part of the teaching task. The teacher identified being respected and accepted as a mastery experience. The teacher also based that success on verbal or written feedback provided to her by students and parents. In 2007, Tschannen-Moran and Hoy examined several potential sources of teacher efficacy beliefs in novice and career teachers. In attempting to quantify mastery experience, they defined a mastery experience as the teacher's perception of success. The definitions of mastery in these studies were not connected to student performance nor were they connected to the best practices involved in teaching.

One study has examined mastery experiences through the lens of its impact on the implementation of a new teaching strategy. Tschannen-Moran and McMaster (2009) studied the impact of professional development on teacher efficacy, conducting a quasi-experimental study testing the potency of different sources of self-efficacy for teachers implementing a new teaching strategy. Teachers from 9 schools participated in 4 different types of professional development in order to implement a new teaching strategy. Findings indicated that the most influential professional development included a mastery experience embedded in the teacher's regular teaching context. In this case, the strategy was implemented in the teacher's classroom with the support of a coach who provided verbal feedback. Teachers receiving this type of professional development increased their self-efficacy for reading instruction, as well as their implementation of the Tucker strategy.

In examining mastery experiences, it may be important to ensure that the mastery experience is defined through a connection to a teaching task that is a best practice of effective instruction. Just as the instruments for measuring teacher efficacy have been connected to rubrics describing effective instructional practice, the lens for qualitative analysis of the sources of efficacy may need to view the teaching tasks for mastery experiences in a similar way.

Vicarious Experiences

A vicarious experience occurs when a teacher experiences another teacher engaging in a mastery experience. Research on vicarious experiences suffers from the same limitation as that of mastery experience, in that the definition of the mastery experiences has been disconnected from best practices in teaching. A few studies, however, have examined vicarious experiences in their relationship to professional development and implementation of new methods of teaching. For example, Bruce and Ross (2008) examined the effects of peer coaching on teacher efficacy in twelve teachers who were engaged in an intensive professional development program that included vicarious experiences in the form of modeling by other teachers. They found that teachers experienced an initial dip in efficacy as they began the program due to a new understanding of how mathematics should be taught. However, by the end of the program teachers reported increased efficacy and were teaching math using the new methods they had been taught.

In the before mentioned study by Tschannen-Moran and McMaster (2009), a decline in teacher efficacy was seen in treatment groups that did not include coaching support for implementation of the strategy. These groups observed other teachers successfully implementing the strategy, but were not supported with coaching in the classroom. It could be that seeing someone successfully implement the strategy and move students that the teachers had previously discounted caused a lowering of self-efficacy. This connects to Wheatley's (2002) hypothesis that teacher efficacy doubts are necessary for the successful implementation of new strategies. However, while these teachers experienced a dip in efficacy that could have led them to a change in practice, the lack of support in implementing the new strategy had the opposite effect, resulting in an overall lowering of efficacy for reading instruction and a less than quality

implementation of the new strategy. Since efficacy was measured after the treatments, there is no way to tell if the other treatment groups also experienced that initial dip in efficacy.

Verbal Persuasion

In many of the previously mentioned studies, verbal persuasion in the form of coaching and/or feedback when in conjunction with another source of efficacy strengthened that source creating positive changes in teacher efficacy beliefs. In both the Tschannen-Moran and McMaster (2009) and the Bruce and Ross (2008) study the addition of verbal feedback and coaching in conjunction with a vicarious experience caused an increase in teacher efficacy. In the Milner (2002) and Milner and Woolfolk Hoy (2003) studies, verbal feedback provided evidence of a mastery experience for those teachers.

In many cases the sources of efficacy work together to form a system for efficacy development. This can be seen in the Chong and Kong (2012) qualitative analysis of how teacher collaboration created the sources of efficacy: mastery experience, vicarious experience, verbal persuasion, and physiological and emotional experiences (affective states). They studied ten teachers from the mathematics, humanities, and science departments of an all girl's school in Singapore. The school utilized a collaborative learning structure for teachers called Lesson Study. Chong and Kong found that the Lesson Study structure provided teachers with opportunities to generate mastery experiences as they planned and delivered lessons collaboratively. This collaboration allowed the teachers to try new strategies with greater confidence. Having a peer observe the implementation of the collaboratively designed lesson, also aided in the ability to understand why some aspects did not go as planned. This also created a vicarious experience for the observer. They were able to anticipate some problems they might encounter, but also felt better about being able to try the new practice themselves after having

seen a peer be successful. Verbal persuasion occurred in the form of constructive feedback and in the encouragement and support provided through the collaborative group work. Finally, they found that teachers' affective states changed throughout the process of implementing the Lesson Study. At first the process provided some anxiety for teachers, but in the end they found the process to be worthwhile.

Affective States

Bandura (1977) defined affective states as emotional arousal that impacts the way a person copes with and manages stress. Many of the studies that have explored the sources of teacher efficacy development have failed to study this source. For example, Tschannen-Moran and Woolfolk Hoy (2007) only explored master experiences and verbal persuasion in their study of the antecedents of self-efficacy beliefs. One study was found that linked the other three sources of efficacy development to their impact on teachers' affective states. As previously mentioned, Chong and Kong (2012) found that teachers' stress and anxiety levels changed as they were engaged in the collaborative learning structure. It could be that the other three sources used to facilitate changes in efficacy do so by helping teachers manage and cope with their stress.

Changing Efficacy

While Bandura (1997) suggested that efficacy is resistant to change once it is established, several research studies suggest that efficacy can be changed. Tschannen-Moran and McMaster (2009) in their professional development study conducted a repeated measures ANOVA which revealed significant increases in teacher-self efficacy over time, regardless of the treatment group [$F(3, 89) = 33.42, p < .01$].

The sources of efficacy can be created through modeling, peer coaching, and new mastery experiences. In the studies where changes were seen, concerted effort was being made to create the opportunities for efficacy development. Evaluation systems may be able to create these same opportunities. In order to understand how efficacy for effective teaching practices is developed, a clear definition for effective teaching must first be clearly outlined. In the studies above, specific aspects of instruction were targeted: new instructional strategies or teaching tasks defined by the teacher themselves. An understanding of how a teacher evaluation system influences teacher efficacy may aid in our ability to better understand how teacher efficacy for effective practice is developed.

Effective Evaluation Systems

There are several factors that are related to the effectiveness of an evaluation system designed to result in changing teacher practice. These factors can be divided into characteristics of the system and characteristics of implementation. The system characteristics refer to the characteristics of the system as it was designed. Modifying the design of the system is the only way to influence these characteristics. Whereas implementation characteristics refer to those characteristics that emerge once the system is being implemented. These characteristics are influenced by the people involved in the implementation, rather than by changes in the system itself.

System Characteristics

The most effective teacher evaluation systems are both formative and summative, providing for both the organizational needs of the school and the individual needs of the teacher (Weber, 1987; Colby, Bradshaw, & Joyner, 2002; Conly & Glasman, 2008; Delvaux, Vanhoof, Tuytens, Vekeman, Devos, & Petegem, 2013). According to Colby et al. (2002), experts state

that evaluation systems should be designed with a close connection between the methods of the system and the goals and purposes of the evaluation. Using an evaluation system for accountability purposes requires one that is “highly evolved” with clearly defined criteria that differentiates proficiency levels and indicates what will serve as evidence for each item (Danielson, 2012). A clear and defined set of criteria for evaluation allows teachers to know how they will be evaluated and provides a guide for evaluators to know what to look for in the lesson. The criteria should be research based and validated so that a high level of performance on the criteria is a predictor of high levels of student learning. In addition to providing clarity to different levels of teacher competency, the criteria should provide a guide for teacher reflection and feedback. This provides the evaluation system with the function of also being utilized formatively to improve instructional practice (Danielson, 2012).

Another characteristic of effective evaluation systems is that they emphasize collaboration between teachers and evaluators (Weber, 1987). Often this collaboration occurs in the form of how feedback is provided to the teacher following the evaluation. The structure for providing feedback should ensure timeliness, specificity, and credibility as much as possible. McLaughlin (1986) suggested that feedback should be provided when the event is fresh in both the mind of the evaluator and the teacher; it should be specific and linked to student data; and it should be credible, coming from a trusted source. Colby, Bradshaw, and Joyner (2002) shared that research suggests effective evaluation systems focus on growth for teacher practice and become a guide for new learning.

Effective evaluation systems also utilize multiple and varied data on teacher performance. This includes multiple observations of performance, as well as multiple data sources such as lesson plans, student achievement data, and/or samples of student work in addition to direct

observation of instruction (Educational Partners Inc., 2007). Colby, Bradshaw, and Joyner (2002) state:

Both the theoretical literature (Beck, 1994; Beerens, 200; Danielson & McGreal, 2000; Ellet, 1997, McConney, 1995; McConney, Schalock & Schalock, 1997; Peterson, 1995, 2000; Stronge, 1997) and empirical research (Dawson & Ackner-Hocevar, 1998; Stiggins & Duke, 1998) advocated the use of multiple data sources as an approach leading to more effective evaluation of teachers. The use of multiple data sources was viewed as offering more advantages than single source data collection (Bradshaw & Glatthorn, 2001; Ellet, 1997; Peterson, 1995, 2000; Scriven, 1988; Stronge, 1997). (p. 4)

Effective evaluation systems are structured to meet the needs of the school and teacher by being both summative and formative, including detailed criteria for measuring teacher performance, including a structure for meaningful feedback and collaboration, and including the use of multiple measures of teacher performance.

Implementation Characteristics

Once a system is chosen, research suggests several factors are key to successful implementation of that evaluation model.

Perception of Fairness

First, evaluators must be trained and competent to use the evaluation system (Weber, 1987). Teachers should also be trained and understand the system that will be used to inform their practice. Shared training clarifies the role of all participants in the evaluation process and creates a shared language around which to discuss instruction (McLaughlin, 1986). In addition, ongoing professional development for quality assurance ensures that the evaluation system scoring remains accurate and reliable (McLaughlin, 1986). Whether or not the teacher views the

evaluator as a credible evaluator impacts the teacher's willingness to accept the feedback provided (Weber, 1987). Delvaux et al. (2013) explain:

If teachers do not consider the evaluator competent, they fear receiving an undeserved assessment (Milanowski & Heneman, 2001). A credible evaluator 1) has the required competency to evaluate teachers, 2) has considerable experience teaching, 3) is familiar with the subject of the teacher, and 4) has enough opportunities to observe and follow-up with the teacher (Chow et al., 2002; Milanowski & Heneman, 2001; Mo et al. 1998). (p. 3)

Ensuring validity and accuracy assist in establishing credibility for the evaluator. These factors are important to maintaining positive perceptions of fairness for teachers. A lack of perceived fairness is often related to negative reactions towards an evaluation system (Delvaux et al., 2013).

Attitude of the Principal

The attitudes of school leaders can also impact the validity and accuracy of the evaluation measures. Kimball and Milanowski (2009) explain:

A school leader who views the performance evaluation system as too much work or just another mandate is likely to spend less time observing teacher behavior and making careful assessments than one who sees performance evaluation as a tool for instructional improvement. (p. 39)

The principal's attitude influences the teacher's understanding of the evaluation's purpose and goals, as well as the teacher's perception of the evaluation as a process. Teachers who view the process as something that will help them improve their teaching are more likely to utilize the system for instructional improvements (Colby et al., 2002).

Positive Relationships

Another characteristic of successful implementation is the relationship between the evaluator and the teacher. Successful evaluation systems function from a shared culture of collaboration and improvement (Weber, 1987). In addition, a feeling of reciprocity, or mutual respect, is also important (Weber, 1987). Bird and Little (1985) further describe the idea of reciprocity as it relates to the evaluator/teacher relationship stating that first the observer must assert the knowledge and skills needed to evaluate the teacher's classroom. Then the teacher must defer to the observers assertion. The observer then must display their assertions, providing evidence from the lesson that makes sense to the teacher and is a valid interpretation of the lesson. Next the teacher should respond to the assertions made by trying something new based on the evidence. Finally, the observer should also be working to continuously improve their practice. These actions create a positive relationship between the observer and the teacher built on a mutual purpose.

High Quality Feedback

The quality of the feedback provided to the teacher is another characteristic that impacts the successful implementation of an evaluation system (Delvaux et al., 2013; Weber, 1987). Danielson (2012) stated that the post conference conversation should invite teachers to “reflect on their practice and strengthen it in ways described by the instructional framework used” (p. 36). This requires the evaluator to be skilled in conducting coaching conversations, as well as in focusing those conversations around potentially meaningful changes to teacher practice. Teachers want feedback that offers specific suggestions for improvement, rather than general statements about the quality of their instruction (Conly & Glasman, 2008). McLaughlin (1986) stated:

Teacher evaluation potentially is such a powerful feedback mechanism because evaluation is a way of giving meaning to activity. But it can play this important role only if it is received, if it is heard, and if it is acted upon. (p. 69)

In addition, when feedback is supported by follow-up in the classroom, more meaningful changes to teacher practice occur.

Tennessee Educator Acceleration Model

The Tennessee Educator Acceleration Model (TEAM) is a teacher evaluation system designed to combine frequent observations, constructive feedback, measures of student learning, and aligned professional development opportunities to create a model of continuous improvement aimed at increasing teacher effectiveness across the state (team-tn.org). TEAM is designed to be both formative and summative, utilizes a defined set of criteria to evaluate teacher performance, includes a structure for providing feedback and initiating teacher collaboration, and uses multiple measures of teacher performance.

The system utilizes a rubric that was developed in conjunction with the National Institute for Excellence in Teaching (NIET) as its criteria for measuring teacher performance. According to the Tennessee Department of Education, the decision was made to partner with NIET to develop this rubric based on NIET's positive field test results, research linking the original rubric to student achievement, NIET's ability to provide high-quality evaluator training, and availability of the NIET Best Practices Portal for additional professional development options for teachers (team-tn.org).

The rubric provides clearly defined levels of performance. The rubric is made up of 3 domains: Environment, Planning, and Instruction. Each domain contains indicators, or major

headings of what makes an effective lesson. Then each indicator contains descriptors. These describe the indicator at each performance level.

The rubric has also been correlated to measures of student performance. The NIET rubric was originally developed in a study by Schacter and Thum (2004). The purpose of the study was to create teaching performance standards and rubrics, rate teacher performance using those standards, and determine the magnitude of student achievement gains attributed to differences in teacher quality based on those standards. To create the rubric, Schacter and Thum (2004) surveyed three bodies of research on teaching. Then the rubric was given to an expert panel consisting of “curriculum specialists, administrators, and teachers to review, comment on, and add to or delete items from [their] standards draft” (p. 412). Following the creation of the rubric, Schacter and Thum trained five graduate students in an extensive process to ensure inter-rater reliability and accurate scoring using the newly created rubrics. The five graduate students then conducted observations of 52 elementary teachers using the rubric. Each teacher was observed 8 times over the course of a 9-month period. Each teacher was observed at various points throughout the day and teaching a variety of subject matter. Two of the observations were scheduled ahead of time; the other six were unannounced. The results of the study indicate that higher quality teaching, as defined by the performance standards and rubrics, was positively correlated to higher student achievement gains.

TEAM utilizes the scores on the evaluation observation as a catalyst for providing feedback and promoting teacher reflection. Following each observation, the evaluation process includes a post conference where teachers are asked to reflect upon an area of reinforcement (a strength) and an area of refinement (something to strengthen). The post conference procedure includes the evaluator identifying the area from the rubric on which to focus teacher reflection,

planning and asking coaching questions to guide the teacher to reflect on that area, providing evidence of why that area was a strength or weakness, and recommendations for continuing to strengthen that area or suggestions for improving the area of weakness. The structural procedure provided can serve as a springboard for additional support or professional development as well.

TEAM also uses multiple measures to establish a teacher's overall effectiveness rating. For TEAM, these multiple measures include multiple observations using the NIET rubric by certified evaluators, as well as the use of student achievement data. The rubric scores make up 50% of the teacher's overall effectiveness rating. The other 50% of a teacher's effectiveness rating is made up of quantitative student data. Student growth data makes up 35% of this 50% and student achievement data makes up the other 15%. In cases where teachers do not have individual classroom data, school-wide data are used.

Two recent studies analyzed teacher and principal perceptions of the TEAM system. Bryant (2013) analyzed principals' perceptions both about the impact of the system on the quality of teaching and perceptions on the principal's ability to implement the system with fidelity. Bryant (2013) found that overall principals perceived the TEAM system as having a positive impact on instruction through reflection and collaboration around teaching practices. However, they struggled with time management and being able to implement the system with fidelity. Moran (2013) examined teachers' perceptions of the influence of TEAM on literacy instruction and examined how context impacted their perceptions. Moran (2013) found that teachers in schools with principals more engaged in the TEAM process had more positive perceptions of the system and made more meaningful changes to their practice. These two findings suggest that the principal's investment in the system, as well as teachers' perceptions of the system, are contributing factors to the evaluation system's ability to impact teacher practice.

This idea is consistent with Colby, Bradshaw, and Joyner's (2002) literature review of teacher evaluation. They found that successful evaluation systems required principals to see the evaluation system as a tool for instructional leadership and teachers to perceive the process as a way to improve their performance.

Structurally, the system is designed around all the elements of an effective evaluation system. However, variances in implementation influence the overall effectiveness the evaluation system can have on improving teacher performance. Despite the variances in the fidelity of implementation and perceptions of the TEAM system across the state, Tennessee experienced large improvements in student achievement according to NAEP in 2012-2013. News reports indicated that the state had the largest academic growth of any state and the largest single testing cycle growth on record (<http://news.tn.gov/node/11644>).

Conclusions

Teacher efficacy has evolved and changed over time. Beginning with the RAND studies in 1975 and evolving through the early 21st century, teacher efficacy has a long history of measuring a teacher's belief in whether or not his/her actions will impact student outcomes. Separating the teacher's belief in their own ability to effectively enact best teacher practices from their belief about teaching and learning, teacher self-efficacy becomes a separate construct and measurable in its own terms. While there is a plethora of research to suggest that teacher efficacy has a positive correlation to student outcomes, little to no research exists connecting the construct of teacher self-efficacy to student achievement. In fact, few researchers have looked at teacher self-efficacy as a construct separate from teacher efficacy.

Tennessee's evaluation system is an organizational structure with the potential to influence both of these efficacy constructs. Designed around the elements of an effective evaluation process, TEAM has the potential to change teacher practice. Utilizing a rubric that focuses on specific teacher practices, the system might be creating opportunities to impact teachers' beliefs about their ability to use those best teaching practices. In addition, with its focus on how those actions impact student outcomes, the system might also be creating opportunities to influence what teachers believe about how their teaching behaviors impact student learning.

Understanding how the structures of TEAM, as well as the implementation of TEAM, impact teacher self-efficacy could be beneficial both for strengthening the system itself and for designing other evaluation systems in the future. To understand the extent to which TEAM acts as an intervening variable to teacher self-efficacy, and to understand how TEAM influences teacher self-efficacy, a mixed methods study will be conducted. Chapter Three will explain the selection of mixed methods as the methodology for this study. The role of the researcher will be explored. Then a description of the data collection procedures will be provided, followed by a description of the correlating data analysis. Finally, an explanation of the methods for verification will be given.

Chapter 3

Methodology

The purpose of this mixed methods study was to examine whether the TEAM evaluation model, as an intervening variable, influences teacher self-efficacy and to explain how the various characteristics of implementation of the TEAM model impact efficacy development. This chapter begins with an overview of the research design as it was intended, as well as adjustments that were made during the IRB process. An explanation of the methods, the justification for their use, and an explanation of the changes made is also provided. Next the role of the researcher is clarified. This is followed by a description of the procedures intended for data collection and a description of the actual methods used for collection. A description of the final procedures for data analysis is then provided. The chapter concludes with an explanation of the methods that were intended for verification, as well as the actual methods used. The University of Tennessee's Institutional Review Board evaluated all of the methods to be utilized in this study, and accepted the final methods used after several reiterations.

Research Design

This study employed a mixed methods design. Creswell (1994) defined four mixed methods designs: sequential studies, parallel/simultaneous studies, equivalent status studies, and dominant-less dominant studies. Tashakkori and Teddlie (1998) added a fifth, designs with multilevel use of approaches. This study was designed to utilize a parallel/simultaneous mixed method design as the qualitative data collection phase and the quantitative data collection phase were intended to occur simultaneously (see Figure 1). However, in order to ensure the safety of participants against retribution for participation in the study and to protect them again undue stress that may have been caused by the observation of their evaluation conferences, data

collection was limited to survey data and interviews only. This changed the collection design to a sequential design in which the quantitative data were collected prior to the qualitative data. (See Figure 2).

In this study the quantitative data analysis and the qualitative data analysis were intended to have equal status: QUAN/QUAL. The quantitative data were to be utilized to answer the first research question: To what extent does the implementation of TEAM act as an intervening variable for teacher self-efficacy? However, limited participation resulted in an inability to conduct some of the analysis intended.

The quantitative data were used to add clarity to the qualitative data which answered the second research question: How does the implementation of TEAM influence teacher self-efficacy? A preliminary analysis of the quantitative data was made prior to the conduction of the interviews in order to finalize the interview protocols.

Rationale for the Design

Tashakkori and Teddlie (1998) stated that the evolution of research methods has resulted in many researchers adopting the tenants of paradigm relativism, using whatever method is appropriate for the study. The selection of methods for this study was determined based on the questions to be explored:

1. To what extent does the implementation of TEAM act as an intervening variable for teacher self-efficacy? (Quan)
2. How does the implementation of TEAM influence teacher self-efficacy? (Qual)

The first question is positioned in the quantitative paradigm, requiring the “testing of a theory

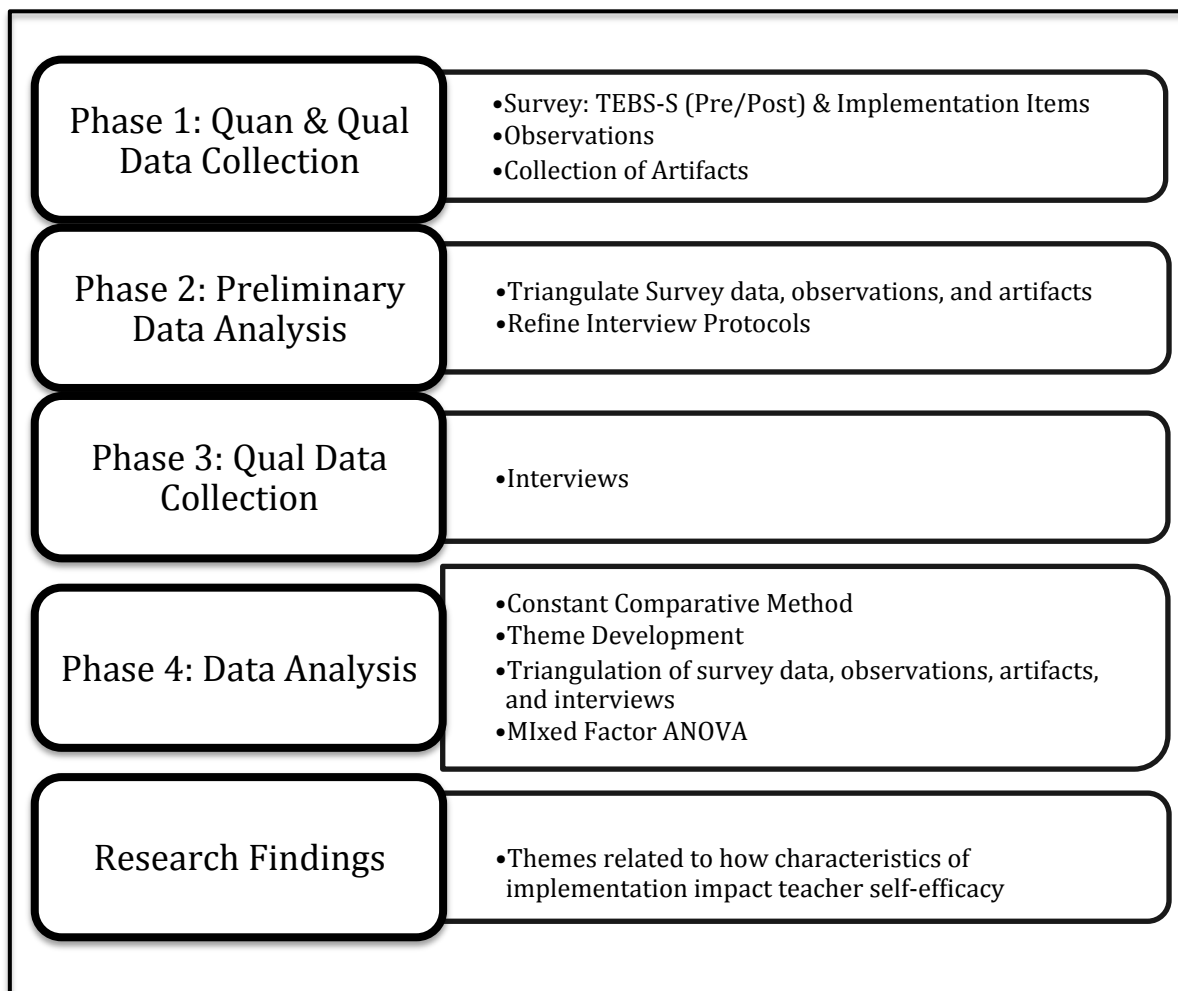


Figure 1. Research Design Flowchart

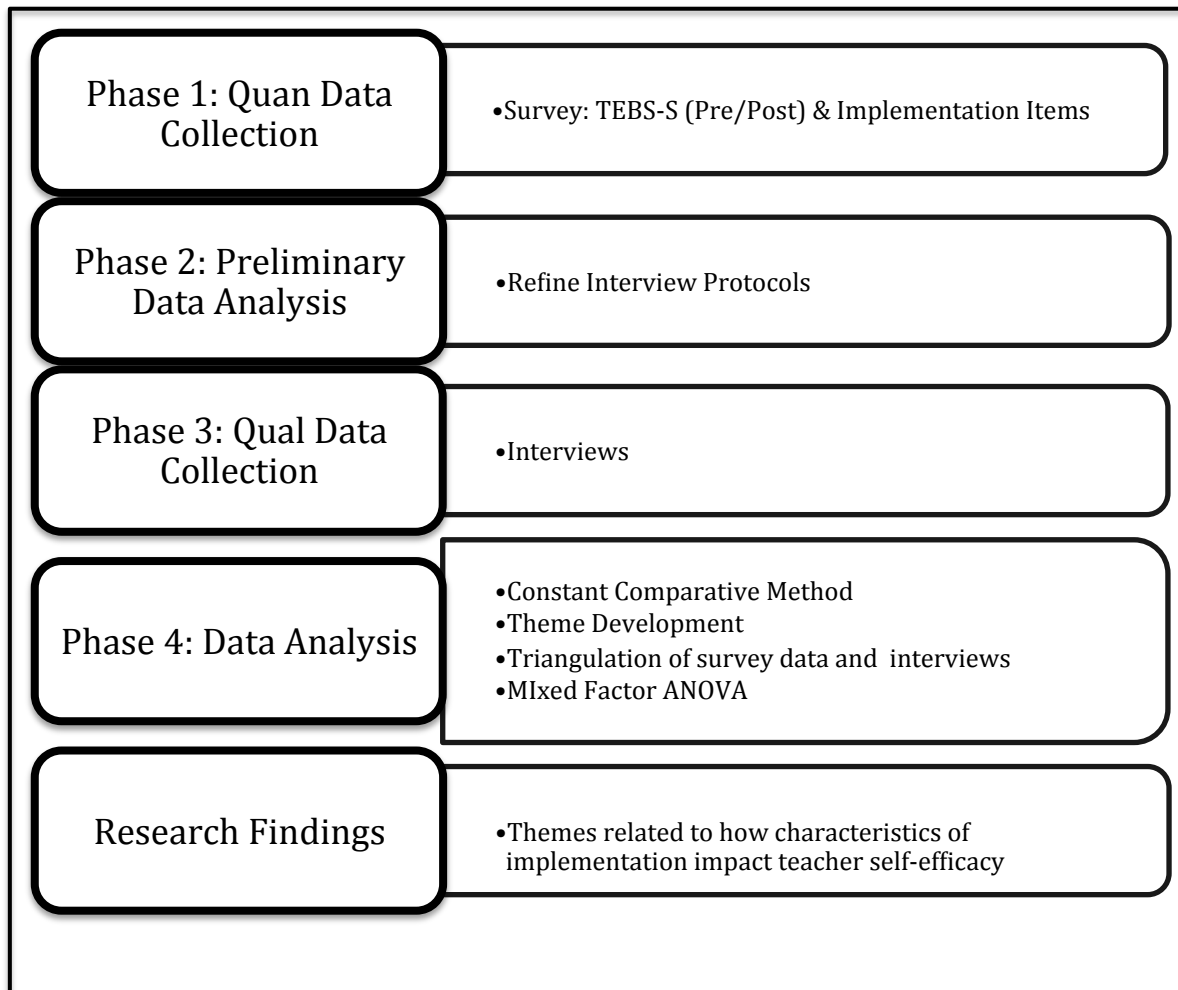


Figure 2. Research Design Flowchart Revised

composed of variables, measured with numbers, and analyzed with statistical procedures, in order to determine whether or not the predictive generalizations of the theory hold true” (Creswell, 1994, p. 2). The quantitative phase of this study can be defined causal-comparative. Causal-comparative research attempts to determine the cause for differences in behavior (Gay, Mills, & Airasian, 2009). In causal-comparative studies, the groups are already formed and are not manipulated by the researcher (Gay, Mills, & Airasian, 2009). Such is the case here. All teachers were being studied within the context of the established set of implementation processes within their school. The independent variable in this study is the level of implementation of TEAM. The dependent variable is the extent of change in teacher self-efficacy. The quantitative portion of this study was intended to determine if differences in implementation of TEAM are connected to variances in changes in teacher self-efficacy.

The second question is positioned in the qualitative paradigm. The qualitative question seeks to understand how efficacy is influenced and requires examining the view of the participants to build a “complex, holistic picture, formed with words” (Creswell, 1994, p. 2). In order to build this picture, data in the form of surveys, observations, artifact collection, and interviews were intended to be collected (See Table 1).

Since the questions were positioned in both paradigms, the method selected for this study was mixed method. Mixed method studies use triangulation techniques to study the same phenomena using both qualitative and quantitative methods (Tashakorri & Teddlie, 1998). Yin (2009) stated, “mixed methods research can permit investigators to address more complicated research questions and collect a richer and stronger array of evidence than can be accomplished by a single method alone” (p. 63). Mixed method designs allow for the researcher

to “work back and forth between inductive and deductive models of thinking” (Creswell, 1994, p. 178). The quantitative aspect of this study added depth and clarity to the qualitative portion.

Table 1.

Research Questions Related to Data Sources

Research Question	Survey Questions (TES; TBES-S)	Interview Questions	Observation	Artifact
1. To what extent does the implementation of TEAM act as an intervening variable for teacher self-efficacy? (Quan)	✓	X	X	X
2. How does the implementation of TEAM influence teacher self-efficacy? (Qual)	✓	✓	✓	✓

Following the IRB process, the collection of observations and artifacts was omitted from the qualitative data collection. The evaluation process can be stressful for many teachers. In addition, interviewing teachers about their experiences during the evaluation process, which requires talking about their principal, makes them extremely vulnerable, especially if they are teachers without rights to due process. Since the observation of the post conference would include both the evaluator and the teacher, adequate measures to protect the teachers’ confidentiality could not be secured. In order to ensure that participants were not put into undue stress as a result of being observed and having their evaluation artifacts collected, and that they

were more adequately provided confidentiality in their interviews; the observations and collection of artifacts were eliminated from the study.

Role of the Researcher

The role of the researcher is influential in qualitative research. Gay, Mills, and Airasian (2009) stated, “it is critical to the success of the study that the researcher establish that he or she will fit comfortably with the participants” (p. 113). To gain entry into the school district, I emailed and had a telephone conversation with the Director of Schools. I communicated my interest in studying the effects of the implementation of the evaluation system on teacher self-efficacy. The Director of Schools agreed that the district would be a good location for such a study and that he would be interested in the findings. In addition, I agreed to provide professional development around the evaluation system following the conclusion of the study. My current position involves training principals and master teachers on how to make the evaluation system effective, and I provide job-embedded professional development utilizing the TAP system in a neighboring county. I have not provided any training to the participants in the sample county.

In order to gain access and establish the purpose of the study, the Director of Schools emailed the principals to inform them that he had agreed to allow me to conduct the study within the district. Each principal was asked to email me if they were interested in participating. Following the IRB process, I contacted each principal again via telephone, explained the study, and asked if they would still be willing to participate. I agreed to provide feedback on the evaluation process to each principal and/or evaluation team at each school site following the conclusion of the study. I then met with each staff during an after school staff meeting. I

explained the study, as well as the potential risks for participation. I also answered any questions that teachers had at that time.

The role of the researcher in the analysis of the data is as an instrument for constructing meaning from the data. A researcher's bias, values, and judgments become an integral part of the meaning constructed in the study (Creswell, 1994). As a trainer and supporter of the TAP system, I believe the evaluation system can be a useful tool in improving teacher self-efficacy. I also believe that the evaluation system is only as powerful as the way in which it is implemented and the belief that is placed in it for changing instruction. My goal in this research is to uncover what needs to be in place for positive changes in teacher self-efficacy to occur and to describe any negative changes so that principals, districts, and state officials can see the need for increased professional development and growth to improve these implementation characteristics of the evaluation system so that it can be effective in all settings. I believe that all children are capable of reaching higher levels of academic achievement than they are currently. I believe that changing teacher practice in a positive way will increase teacher effectiveness and have a direct impact on improving student outcomes.

To make my decisions during the analysis clear to readers, a code map was utilized (See Table 4). Anfara, Brown, and Mangione (2002) suggested that "making all aspects of the analysis process open to public inspection" creates an "audit trail" that increases the dependability and reliability of the findings (p. 31). In addition, it was intended that data would be triangulated between observations, interviews, artifacts, and quantitative analysis. However, the deletion of the collection of observation data and artifacts limits the ability to triangulate the data. Finally, member checks were utilized to ensure that an accurate picture of the participants

experience had been captured and explained. This process is explained in detail in the section: Methods of Verification.

Site and Participants

The goal of the sample selection was to obtain a sample that would be “information-rich” representing a multitude of implementation characteristics and a variety of teacher experiences (Merriam, 2009, p. 77). To obtain an information rich sample, purposive sampling was used. Purposive sampling is “the process of selecting a sample that is believed to be representative of a given population” (Gay, Mills, & Airasian, 2009, p. 134). The decision was made to narrow to one district so that all of the participants would be operating under the same evaluation policies with access to the same training and professional development. Furthermore, the district’s Director of Schools self-identified the district as having a variety of levels of TEAM implementation. The selection of this district allowed for maximum variation in the characteristics of implementation while minimizing the variation in district resources. Merriam (2009) refers to this type of purposive sampling as maximum variation sampling, “seeking those who represent the widest possible range of characteristics of interest for the study” (p. 79). In this case the characteristic of interest is the implementation of the evaluation process in each school.

The study was then delimited to elementary school teachers in the district. Elementary teachers were chosen as the sample frame because the daily function of an elementary classroom teacher varies greatly from that of a middle school, high school, or special area teacher. Elementary teachers typically teach the same students throughout the day and teach multiple contents. Whereas middle and high school teachers teach multiple groups of students and generally focus on one content. According to Tschannen-Moran and Woolfolk Hoy (2001), context plays a role in teacher efficacy. Part of context is the students who are taught.

Elementary teachers experience only one context around which their efficacy is developed, whereas middle and high school teachers may experience multiple contexts.

The site for this study was a rural Tennessee district that consists of eight schools. Five of the eight schools include elementary grades, PreK-5. Three of these five schools are PreK-12 schools. Approximately 97 teachers make up the staff for the elementary grades in the district. Of the five schools with elementary teachers, one utilizes a teacher support model to accompany its evaluation process. This school is using The System for Student and Teacher Advancement (TAP) to add structures for career advancement, job-embedded professional development, and performance pay to accompany the evaluation model. The other schools use the TEAM evaluation model. Three of these five schools agreed to participate in the study.

Quantitative Sampling

No additional sampling technique was utilized for the quantitative data collection. All elementary teachers in the three participating schools were asked to participate in this phase of the study. A statement at the beginning of each survey asked participants to give informed consent.

Qualitative Sampling

Gay, Mills, and Airasian (2009) stated that there are no “hard and fast rules” for determining the number of participants for a qualitative study (p. 136). Merriam (2009) agreed, saying, “What is needed is an adequate number of participants, sites, or activities to answer the question being posed” (p. 80). In the original design, two school sites utilizing the TEAM evaluation model were to be selected for the qualitative portion. Due to differences in the requirements for implementation between TEAM and TAP schools, the qualitative sample was intended to be delimited to only TEAM schools. In TAP schools, all teachers are required to

participate in 4 evaluations that cover all 3 domains on the NIET rubric. In the TEAM schools, the number of evaluations varies based on set teacher criteria. The criteria are the same across all the TEAM schools. In addition, all 3 domains are not assessed simultaneously in all evaluations within the TEAM schools, as they are in the TAP school. However, due to a reduction in the number of schools that agreed to participate and changes in the sampling techniques for the teacher interviews, all three schools were used in the qualitative sample. This will be explained further in this section.

In order to explore the differences in implementation of the TEAM model and how those differences influence efficacy, maximum variation sampling was to be utilized. The sites were to be chosen following analysis of principal interviews using a characteristics of implementation decision matrix. All four TEAM principals were to be interviewed and their responses were to be used to determine a score for each characteristic of implementation. A high score, 5, would have indicated that the responses suggested that the characteristic of implementation was highly evident. A low score, 1, would have indicated that the responses suggested that the characteristic was not evident. Anfara, Brown, and Mangione (2002), explained that every attempt by the qualitative researcher should be made to make their process visible to the reader. The use of a decision matrix, was chosen for this stage in the research because the characteristics for selection were already known (See Table 2). In addition, teachers were asked to complete a survey on the implementation characteristics. The survey scores were then to be averaged to generate a mean score for teachers at each site. The score from the principal interviews and the mean score from the teacher surveys was then to be averaged to generate a final score for each site. The site with the highest score and the lowest score for effective implementation was then going to be selected and invited to participate in the next phase of the study.

Table 2.

Implementation Characteristics Matrix

Implementation Characteristics	Interview Questions that Address	Highly Evident		Not Evident		
<u>Perceptions of Fairness:</u> Teachers view the evaluation process as fair and their scores as accurate.	4,5,6	5	4	3	2	1
<u>Attitude of the Principal:</u> Principal views the evaluation process as a tool for instructional leadership.	1,2,3,9	5	4	3	2	1
<u>Positive Relationships:</u> The relationship between observer and teacher is one of mutual respect and common purpose.	2,4,9	5	4	3	2	1
<u>High Quality Feedback:</u> Feedback is focused on meaningful changes to the teacher's practice.	2,7,8	5	4	3	2	1

However, working with the IRB committee to ensure the safety of participants, resulted in the study changing from fall semester to spring semester. This change in timing, resulted in two schools withdrawing from the study, citing that they had asked their teachers to participate in several other surveys already within the school year. In addition, to ensure that teachers were adequately protected against coercion, teachers were asked to volunteer to be a part of the qualitative sample. Only three teachers volunteered for this portion of the study, one from each site. Therefore, all three schools were used within the qualitative portion and the three teachers who volunteered were used as the teacher interview sample.

It was intended that once sites were chosen, 5 teachers from each school would be selected using random sampling with replacement. Merriam (2009) suggested offering a tentative, approximate number of participants that is adjusted in order to reach saturation or redundancy. The initial selection of 5 teachers per site was to be utilized because ten teachers represented approximately 10% of the total population of elementary teachers in the district. Using random sampling with replacement would have allowed for generalization from the sample to the population (Patton, 1990). The goal was for this group of teachers to represent a typical sample of teachers engaged in the evaluation process at each site. The names of all elementary teachers at each site were to be placed in a container. Participants were to be selected by drawing names from the container. After each name was drawn, the teacher would then have been asked if he/she would be willing to participate in the qualitative portion of the study. The teacher's name would then have been placed back in the container prior to the next name being drawn ensuring that each participant had an equal chance of being drawn (Gravetter & Wallnau, 2011). This process was to be repeated until 5 teachers from each site had been selected. At each site, the evaluator for each participating teacher would also have been asked to participate in the

observation process. If the evaluator was not willing to participate, the teacher would not have been selected and another name would have been drawn.

All interview and observation participants, both teachers and evaluators, would have been asked to give informed consent, sign an additional statement indicating their willingness to participate in the study. Prior to requesting these additional signatures, it would have been explained to each interview and observation participant that confidentiality, but not anonymity would be provided because of the necessity for connecting the various data together for analysis. The observation data, artifacts, interview data, and survey data would all have been connected for triangulation.

Instead, at the conclusion of the first survey administration, teachers were asked to volunteer for the interview portion of the study. If they were willing to participate, they were to check a box and include additional contact information. They were asked to provide a phone number and email address by which they would prefer to be contacted. It was explained that interviews would occur at the location of the participants choosing. All volunteers signed consent forms indicating their willingness to participate in the interview portion of the study, and expressed their understanding that confidentiality, but not anonymity would be ensured. The volunteers were also informed that they would have an opportunity to review the transcript and ask that any portions of their statements be deleted or edited in order to provide an additional layer of protection. In addition, no demographic data were collected in order to add an additional layer of protection for participants.

Data Collection Procedures

Quantitative Data Collection

At the beginning of the evaluation cycle, all teachers were emailed a link to the Teacher Efficacy Belief Scale-Self (TEBS-S) and the Implementation Characteristics Survey and surveys were taken electronically using Qualtrics. Gay, Mills, and Airisian (2009) suggested a response rate of 60% or less can impact the generalizability of results. So, when only four out of the 53 teachers had responded via email, the surveys were re-administered using a paper and pencil version at each school site. In addition a gift card of \$5 was offered to each teacher for participation in the survey portion. This resulted in a response rate of 36%. At the end of the evaluation cycle, the paper and pencil version of the TEBS-self was re-administered to those teachers who had chosen to participate in the study. At this time 8 participants did not continue in the study, resulting in a final response rate of 20.7%.

All the survey data were entered into an excel spreadsheet and then imported into SPSS for later analysis. Procedures for analysis of the data are described later in this chapter.

Instrumentation.

TEBS-self. The TEBS-self was selected as the instrument for measuring teacher self-efficacy for several reasons. First, the instrument is designed to measure teacher self-efficacy beliefs about “tasks that are associated with correlates of best teaching practice” (Dellinger, et al., 2008, p. 756). The language within the measure also resembles the language in the TEAM instructional rubric, the tool used in the evaluation process.

The TEBS-Self was developed by colleagues at Louisiana State University utilizing Bandura’s Guide for Constructing Efficacy Scales. It was then utilized in a pilot study and administered to 470 teachers. The scale was then revised and the number of items reduced from

70 to 51. A second analysis was conducted through an Expert Survey Questionnaire in order to determine the validity of the measure. A group of 46 educators, including teachers, principals, central office personnel, state department personnel, and university faculty, rated each item on the importance of the task to be included. Items were then deleted if they did not perform well or were found to be redundant, reducing the total number of items to 30 (Olivier, 2001).

The TEBS-Self was then tested for reliability in “three independent studies using the TEBS-self with large samples ($n = 2373$ total) of K-6 elementary teachers” (Dellinger, et al., 2008, p. 756).

The instrument consists of 30 questions. Teachers are asked to rank their beliefs in their ability to engage in various teaching tasks on a Likert scale from 1 (weak beliefs in my capability to) to 4 (very strong beliefs in my capability to). For the 30-item one factor solution to the TEBS-S ($n=1437$) the Standardized Cronbach Alpha Reliability Coefficient is .95. The measure can be divided into five subscales: Communication/Clarification (.87), Management/Climate (.85), Accommodation of Individual Differences (.85), Motivation of Students (.78), and Higher Order Thinking Skills (.86).

Characteristics of Implementation Survey. Ten items relating to the characteristics of implementation were developed to aid in determining the level of implementation at each site. Each item was based on the research on evaluation implementation. Information gathered from this survey was designed to provide additional description of each site. This survey was not tested for validity or reliability.

Qualitative Data Collection

Observations. One post conference observation for each teacher in the qualitative sample was intended to be conducted by the researcher. The post conference was chosen for observation

because it is the time designated for feedback to be provided and the only time during the evaluations process where the evaluator engages with the teacher.

During the observation of the post conference, the evaluator was to be observed as the source of efficacy and the teacher was to be observed for how s/he responded to the source. The post conferences were to be scripted for completeness. Both verbal and non-verbal communication was to be recorded, including noting shifts in body language and expressions. The scripts were to be inserted verbatim into a table within a word document with the script of the evaluator on the left and the teacher responses on the right.

As an extension, when the evaluator in the post conference scheduled a follow-up session with the teacher to provide additional support, then that follow-up session was also to be observed and scripted. To ensure that I did not influence the process the evaluator regularly engaged in during the post conference, the teacher and the evaluator were not to be asked about scheduling follow-up visits. If the evaluator scheduled one, then I would have asked for permission to also observe the follow-up session.

In these cases, the evaluator was to be observed as the source of efficacy and the teacher was to be observed for how s/he responded to that source. The follow-up session was to be scripted for both verbal and non-verbal interactions, including noting shifts in body language and expressions. Once again, the script was to be inserted verbatim into a table within a word document with the script of the evaluator on the left and the teacher responses on the right.

However, in order to protect participants against undue stress and provide more confidentiality to their interviews, observations of the post conference were deleted from the study.

Artifacts. Artifacts were to be collected as potential sources of efficacy. All artifacts provided to the teacher from the evaluator were to be collected. These included scripting notes, evidence notes, score sheets, and/or post conference plans.

Once again, to protect teachers from undue stress and to add to the confidentiality of interview responses, these artifacts were not collected.

Interviews. Following the administration of the post survey, the interview protocol was refined. The interview protocol was semi-structured. A defined set of questions was predetermined, crafted to capture the teacher's experience for how various sources of efficacy created in the evaluation process influenced their own beliefs about their ability to implement various teaching tasks. The protocol was designed using several question types: experience and behavior, opinion and values, feeling, knowledge, sensory, and demographic, as suggested by Patton (2002) and referenced by Merriam (2009) (See Table 3).

Refining the Protocol. Several of the originally drafted questions included blanks to insert data from either the qualitative or the quantitative data collection process. So, before interviews were conducted, a preliminary analysis of the quantitative data was to be conducted, followed by a preliminary analysis of the qualitative data. As can be seen in Table 4, the majority of the questions were to be refined using the preliminary data analysis from both the quantitative and the qualitative data.

However, due to the removal of the observations, the interview protocol was adjusted to account for only information that could be drawn from the quantitative data. Due to the fact that I would not be able to observe any of the implementation pieces for myself, the new protocol was readjusted to capture the elements of the evaluation implementation. The protocol was again designed using several question types: experience and behavior, opinion and values, feeling,

knowledge, sensory, and demographic, as suggested by Patton (2002) and referenced by Merriam (2009) (See Table 5). Only questions five and six utilized information from the quantitative data (See appendix for final interview questions).

Preliminary Analysis of Quantitative Data. For the preliminary quantitative analysis, the pre and post survey data for each teacher participating in the qualitative analysis was loaded into a table (See Table 6). The 2 items demonstrating the largest changes for each teacher were then utilized as examples for questions 5 and 6 in the interview protocol. If both positive and negative changes occurred for the same teacher, then the largest positive and the largest negative change were selected.

Each teacher was interviewed individually following the completion of the evaluation cycle. Interviews were scheduled for a time that was convenient for each teacher and were conducted at a location of their choice. Interviews with each teacher were taped using a digital recorder. Each teacher's survey code for the quantitative survey data was read into the recorder prior to the beginning of taping.

The interviews were then transcribed by the researcher using the listen and type method. This method was used to ensure accuracy of the transcripts (Johnson, 2011). Once the transcriptions were complete, they were sent to the participants for them to review and edit as they saw fit. The documents were then uploaded into ATLAS.ti for analysis.

Data Analysis

Quantitative Analysis

The survey data from administration of the TEBS-Self were input into SPSS for analysis. A mixed-factor ANOVA (analysis of variance) was intended to be conducted. The mixed-factor

ANOVA was selected because there is one within-subjects factor, evaluation time (pre- / post-) and one between-subjects factor, school site. An analysis of the within-subjects factor was conducted to analyze whether the elementary teachers as a group had a statistically significant change in their teacher self-efficacy over the course of the evaluation cycle. Next, an analysis of the between-subjects factor was to be conducted. This ANOVA would have analyzed differences between each school site. For both, the dependent variable would have been the difference between teacher self-efficacy scores from pre- to post- evaluation times. All teachers surveyed were to be included in this analysis. However, the low response rate made this analysis impossible.

A follow-up contrast was also to be conducted comparing the mean change in the TAP school to the mean change in the TEAM schools to determine if there was a significant difference between the change seen in the TAP school versus the change seen in the TEAM schools in the district. Once again, the low response rate made this analysis impossible. Instead, an analysis of the differences between the efficacy levels of the three schools was conducted using the pre-test means to determine if there was a significant difference between the initial teacher efficacy levels of the three different school sites. In addition, the mean changes for each site were provided.

Table 3.

Questions Defined by Source of Efficacy and by Question Type

Question Number	Source of Efficacy	Question Type(s)
T1	Dependent upon teacher's experience	Experience and Behavior
T2	Verbal Persuasion Affective States	Feeling
T3	Verbal Persuasion Vicarious Experience	Experience and Behavior
T4	Verbal Persuasion	Experience and Behavior/ Opinion and Values
T5	Dependent on teacher's experience	Opinion and Values/ Sensory
T6	Dependent on teacher's experience	Sensory
T7	Dependent on teacher's experience	Opinion and Values
T8	Affective States	Feeling
T9	Dependent on teacher's experience	Feeling/ Sensory
T10	Dependent on teacher's experience	Feeling/ Sensory

Table 4.

Question Refinement by Data Source

Question Number	Data Source Used for Refinement		
	Observation of Post Conference	Quantitative Analysis by Teacher	Observation of Follow-up Support
T2	✓	X	X
T3	✓	X	X
T4	✓	X	X
T5	X	✓	X
T6	X	✓	X
T7	X	X	✓
T8	X	X	✓
T9	✓	X	✓
T10	✓	X	✓

Table 5.

Question Defined by Evaluation Implementation Characteristic and Question Type

Question Number	Characteristics of Evaluation	Question Type(s)
T1	All	Experience and Behavior
T2	Quality of Feedback	Experience and Behavior
T3	Quality of Feedback Potential for Perceptions of Fairness, Relationship to Evaluator, and Principal Attitude	Opinion and Values
T4	All	Experience and Behavior/ Opinion and Values
T5	Dependent on Answer	Opinion and Values/ Sensory
T6	Dependent on Answer	Sensory
T7	Relationship with Evaluator	Opinion and Values

Table 6.

Change in Teacher Efficacy for Qualitative Sample

Teacher	Pre	Post	Change	Items with Positive Change	Items with Negative Change
Teacher 1	3.63	3.9	+.27	5, 10, 16, 20, 21, 22, 23, 24, 27	29
Teacher 2	3.07	2.83	-.24	28, 22	2, 8, 11, 14, 18, 19*, 21, 25
Teacher 3	3.7	3.77	+ .07	2, 11, 12, 14,	6, 25

*Indicates question with greatest amount of change

Due to the limited number of participants, the transcripts for each teacher were then analyzed to attempt to identify any trends in the data. Examining the amount of change for different teachers and the items where any changes occurred allowed for trends to be identified. These will be explained in the findings section.

Qualitative Analysis

A priori codes. A priori codes were established prior to data collection and analysis. These codes were generated from the review of literature. These codes included the four sources of efficacy and the four characteristics of successful implementation of an evaluation system.

First, all of the principal interviews were coded for only the implementation characteristics. These interviews were used in conjunction with the teacher surveys on implementation to provide a description of the implementation at each school site.

All the teacher interviews were then coded in ATLAS.ti by the source of efficacy: mastery experience, vicarious experience, verbal persuasion, and affective states. Then the interviews were coded by the implementation characteristics being explored: perceptions of fairness, attitude of the principal, positive relationships, and quality of feedback.

Open and Analytical Coding. Once the teacher interviews had been coded using a priori codes, the process of open coding was begun.

Within ATLAS.ti, each interview was read and coded using open coding. Before the reading of each interview, it was noted whether teacher self-efficacy had been increased, decreased, or remained stagnant. This was then used as a lens for analyzing the qualitative data. Any segment of data that was deemed beneficial in explaining “how” the teacher’s efficacy was changed by the varying implementation characteristics was coded. Codes were generated from exact words of the participant or from the researcher’s own words. This process was repeated for each interview transcript. Throughout the process of open coding, whenever codes could be combined to form categories, those codes were merged under a new code. Each time codes were merged into a category, a memo was added explaining why the codes were merged. This merging of codes into categories is known as analytical coding (Merriam, 2009). The process of open coding and analytical coding was repeated for each interview until all of the interviews had been coded in this manner.

Development of Themes. The final stage of the qualitative analysis was the development of themes. The categories generated in the analytical coding were combined to form themes that could be used for answering the research question: How does the implementation of TEAM influence teacher self-efficacy?

Methods of Verification

Tashakkori and Teddlie (1998) and Merriam (2009) make reference to Lincoln and Guba's (1985) call for qualitative researchers to ensure that their findings are trustworthy and credible. Anfara, Brown, and Mangione (2008) add to that call that the process through which the researcher comes to conclusions should be made visible to the reader, stating, "good naturalistic inquiry shows the hand and opens the mind of the investigator to his or her reader" (p. 29). To make the decision making process visible, as well as ensure the credibility and trustworthiness of the results, several strategies for verification were employed.

Data Saturation

Merriam (2009) explains that "the best rule of thumb" for ensuring that your findings are trustworthy is that "the data and emerging findings must feel saturated; that is, you begin to see and hear the same things over and over again, and no new information surfaces as you collect more data" (p. 219). In order to ensure that saturation had been achieved, the research began with the aim of a qualitative sample that represented 10% of the total population. However, the limited volunteer pool makes data saturation difficult to reach. Therefore, the only findings and themes that were established are ones that were present in all of the teacher interviews (See Table 7). For the themes presented, saturation was reached through the presentation of the same idea in repeated contexts. So while saturation may not have been possible by hearing the same idea repeated by multiple people, it was reached by hearing the same idea multiple times.

Reflexivity

Throughout the data collection and analysis process, a reflexivity journal was kept by the researcher. Information about what the researcher was thinking throughout the process as well as all decisions made related to the method and process were recorded. In addition, a code map was

Table 7.

Code Map

Fourth Iteration: (Themes-Answers to Questions)			
Question #2: How does the implementation of TEAM influence teacher self-efficacy?			
Feedback connected to students generates efficacy	Giving Power to the Sources of Efficacy	Support beyond the evaluation process itself is key	
Third Iteration: (Categories)			
Did know my kids Didn't know my kids Didn't teach this grade Give student examples Give you research Give you suggestions How it effected my children	Set up for change Teacher trait Perfect score Didn't ask you to change Disagree Not judged Want us to fix it	Need professional development Like to have help Show me a different way There's coaching Want us to fix it Provide support Trial and error	
Second Iteration: (Implementation Codes)			
High Quality Feedback	Attitude of the Principal	Perceptions of Fairness	Relationships
First Iteration: (Efficacy Codes)			
Mastery Experience	Vicarious Experience	Verbal Persuasion	Affective States
Data: Interviews			

Note. Adapted from “Qualitative Analysis on Stage: Making the Research Process More Public” by V. A. Anfara, Jr., K. M. Brown, and T. L. Mangione, *Educational Researcher*, 31(7). P. 32 Copyright 2002 by American Educational Research Association. Used by permission.

generated to demonstrate the progression of the initial codes to categories and then how the categories were developed into themes (See Table 7).

Member Checks

Once themes were established, the researcher met with participants from the study and shared the findings. The researcher asked each participant if he/she found the findings to be plausible and an accurate representation of their experience. Tashakkori and Teddlie (1998) state, “if the informants agree with the interpretations of investigators, then this provides evidence of the credibility of the results” (p. 92). Each participant’s responses to the findings were recorded.

Thick Description

The qualitative findings were described using thick descriptions that referenced all data sources. These descriptions of the findings can be found in Chapter 4. The goal for these descriptions was that they would provide “enough description to contextualize the study such that readers will be able to determine the extent to which their situations match the research context” (Merriam, 2009, p. 229). For this reason, thick descriptions of the implementation characteristics at each school are used to frame the findings.

Triangulation

It was intended that the data would also be triangulated. Triangulation of data refers to ensuring that the findings hold true for all the available data sources. By engaging in a reciprocal process of analyzing findings from interviews by coding for those findings in the observations and artifacts, findings could be corroborated with evidence from multiple sources. According to Creswell (2002), drawing on information from multiple sources increases the accuracy and credibility of findings. However, the omission of data collection from observations and artifacts makes triangulation impossible.

Conclusion

A mixed method approach was chosen for this study based on the questions to be explored. The quantitative portion consisted of an analysis of the mean changes in teacher self-efficacy in order to describe variances in changes to teacher self-efficacy related to implementation of TEAM. The qualitative portion sought to explain how those variances were enacted in the implementation of TEAM through the use of teacher interviews. The results of these methods are explained in the following chapter and a discussion of what those results mean for teachers, evaluators, districts, policy makers, and future research can be found in Chapter 5.

Chapter 4

Analysis and Findings

The purpose of this mixed methods study was to examine whether the TEAM evaluation model, as an intervening variable, influences teacher self-efficacy and to explore how the various characteristics of implementation of the TEAM model impact efficacy development. This chapter will attempt to answer the following questions:

1. To what extent does implementation of TEAM act as an intervening variable for teacher self-efficacy? (QUAN)
2. How do differences in the implementation of TEAM influence teacher self-efficacy? (QUAL)

The chapter begins with a description of the implementation at three different schools. These descriptions will provide a frame for the data analysis that will follow. First there will be an analysis and explanation of the quantitative data. This will be followed by a presentation of themes related to how the implementation characteristics function as sources of efficacy and how they work to influence teacher self-efficacy development.

District Context

The rural school district in this study is situated in the southeastern United States. With a population of a little over 20,000, it is a close-knit community where everyone seems to know each other. The community is rich in heritage, history, and scenic beauty.

The school district is comprised of 8 schools with 220 teachers and around 3,300 students. The student body is 97% white, and 65.3% of the students are economically disadvantaged. In 2014, 30.7% of students in grades 3-8 were proficient or advanced in math and 40% were proficient or advanced in reading. In the 2015 school year, 36.4 % of students in

grades 3-8 were proficient or advanced in math and 34.1% were proficient or advanced in reading. The district also made all 1s on the Tennessee Value-Added Assessment System (TVAAS) for TCAP (Tennessee Comprehensive Assessment Program), the lowest possible score on this student growth measure on a scale of 1 to 5 in both 2014 and 2015. This is a district working to improve outcomes for students.

Three schools within the district participated in the study. Two are PreK-12 schools and one is PreK-5. School A is an elementary school that comprises grades PreK-5 and has a total enrollment of around 700 students. In 2015, the school scored a level 1 composite on TCAP TVAAS, the lowest score for the student growth measure, indicating significant evidence that the school's students made less progress than the growth standard. School B is a Prek-12 school with a student enrollment of nearly 600 students. In 2015, the school scored a level 2 overall composite on TCAP TVAAS, indicating approaching average effectiveness, or that there was moderate evidence that the school's students made less progress than the growth standard. School C is a PreK-12 school with a student enrollment of a little over 500 students. In 2015, the school scored a Level 3 overall composite on TCAP TVAAS, indicating average effectiveness, or that there was evidence that the school's students made progress similar to the growth standard. Each school utilized the evaluation system in their building in varying ways. Below is a description of each school's implementation as characterized by the principal and teachers. These descriptions are intended to frame the data analysis that will be provided later in this chapter.

Characteristics of Implementation Context

The review of literature posited that there are four characteristics of evaluation implementation that sway the effectiveness of the system as a tool for teacher growth: the attitude of the principal, the teacher's perceptions of the fairness of the system, the relationships

of the teachers and the evaluators, and the quality of the feedback that is provided to the teachers. This study examined data to explore the extent to which these implementation characteristics influence teacher efficacy, and then how they do so. Providing context for what the implementation characteristics are at each site is necessary for interpreting the meaning of this data. The following descriptions utilize principal interviews and a teacher survey of the implementation characteristics to provide a context for the four characteristics of implementation discussed in the literature review. (Frequency counts for the survey results can be found in the appendix.) This context will provide a framework for the analysis of the quantitative data that will follow.

School A

Principal Attitude. Colby et al. (2002) explained that teachers who view the evaluation process as something that will help them improve their teaching are more likely to utilize that system for instructional improvement. The principal's attitude and how he views and communicates the purpose of the system can directly impact how teachers understand the evaluation system's purpose and goals. The principal at School A described the evaluation process as a good tool to promote teacher dialogue and reflection. The principal explained, before the TEAM model, "administrators didn't have a tool to discuss instruction other than just the normal twice every ten year observation forms, and so moving to the TEAM model has really allowed administrators more of a hands on purpose or a hands on impact into the instructional process." This view of the evaluation process as a tool to improve practice is reflected further in his description of the role of evaluation in his school. "I really see two things going on. One is to improve instructional practice. It's a feedback loop, more so than it used to be. Secondly, I see it as a bridge to begin a dialogue about deeper issues with instructional practice." Teachers also

reported on the survey that the principal views the evaluation process as a way to help them improve their performance ($M=3.43$, $n=7$). The principal's view of the system as a tool to improve performance carries over into the use of the data from the evaluation process as a driver for professional development in the building.

Perceptions of Fairness. Teachers' perceptions of fairness influence their willingness to accept feedback (Weber, 1987). This includes whether or not teachers believe the evaluation to be valid and the extent to which they believe their observers are credible. On the teacher survey, all the teachers surveyed ($n=7$) agreed that their evaluators were qualified to accurately assess their teaching, however, only 4 out of 7 teachers at School A agreed that their evaluation scores were an accurate reflection of their teaching ability. McLaughlin (1986) suggested perceptions of fairness could be influenced by the work that the administrative team puts into understanding the rubric and ensuring accuracy to their scoring. "My assistant and I, we do 95% of the observations in our building," explained the principal. "So we have a discussion sometimes too, to make sure that we are seeing the same things, and so that way we kind of norm the observations that we are doing. We will especially get into CODE and start looking at comparing what she saw to what I saw." The principal also described additional norming practices he had done with a previous administrator in the prior year.

In addition to these discussions about different indicators, the principal also labeled the connection of the evaluation process to professional development saying, "It's really a nice tool to allow us to differentiate what we are doing in professional development for our teachers." McLaughlin (1986) proposed training centered on the evaluation process for teachers, as well as evaluators, was a key to clarifying the role of all participants. However, 4 out of 7 of the teachers surveyed still perceive the rubric to be an ineffective evaluation tool. This could be because

teachers do not see the connection between the rubric and their student achievement data. The principal expounded on this saying,

We've not done a lot of work to show that validity. We've done a lot of nuts and bolts work on the rubric...here's what we are looking for. I don't know that we've gone to that next step yet to say, 'ok, here are the results.' But I think we need to talk to teachers about that validity. You know, level 5 teachers on scores, here's what we're seeing. Level 1 teachers on scores, here's what we're seeing. Show them that the process does work. I think we've done it on an individual basis, but I don't think we've done it school wide or district wide like we need to.

Positive Relationships. Weber (1987) surmised that successful evaluation systems function from a shared culture of collaboration and create a feeling of reciprocity, where teachers and evaluators share the responsibility for instructional improvements. The first step in generating this sense of shared ownership is the building of trust. Trust is the word that the principal at School A used to describe the relationship between the evaluator and the teacher. "For too long teachers have not trusted administrators [to] have their best interest at heart. They've seen [evaluation] as a competition maybe or a gotcha game or whatever it is, but in order to move forward there has to be a tremendous amount of trust between administrators and teachers." The principal at School A takes great pride in ensuring that the evaluation process results in growth for teachers and the majority of teachers surveyed agreed that they use the suggestions provided to grow professionally (M=3, n=7). Suggestions are provided at the end of each post conference. In describing the post conference process, the principal stated that a successful post conference is when "...the teacher can walk out with a specific idea or hopefully a couple of ideas to address a key component in their instruction that we need to strengthen, and

then they go out and they implement that.” To ensure that the teachers receive suggestions that are useful, the principal at School A spends a significant amount of time searching through resources. This commitment to a quality suggestion demonstrates the principal’s view of his role in the process as being a support to teacher improvement. The teachers surveyed for the most part agreed that the evaluators and teachers work together to improve instruction in their school, with four out of six agreeing or strongly agreeing. In addition, all but one of the teachers agreed that the relationship between evaluators and teachers was a positive one.

High Quality Feedback. The principal’s commitment to providing high quality suggestions is a part of the generation of a structure for high quality feedback on instruction in the building. Most of the teachers surveyed agreed that the feedback provided to them as a part of the evaluation process was meaningful and that their principal utilized the rubric to provide them with frequent feedback on their performance. The principal explained how the evaluation process provides him with a tool for providing feedback saying:

Well, before in the old system...professional educators were evaluated twice every ten years...administrators didn’t have a tool to discuss the instruction other than just he normal twice every ten-year observation forms. So moving to the TEAM model has really allowed administrators more of a hands-on purpose or a hands-on impact into the instructional process.

The formal feedback provided during the post conference is in the form of a reinforcement, something that went well, and a refinement, something to be improved upon. The principal described the post conference process saying, “One you’ve got to share what you see that’s very good. Then you’ve got to say, ‘Here are a few of the things that I think we’ve got to work on.’ I

think that's one of the ways that we've made it more reflective and less of a gotcha kind of game."

To prepare for that post conference, the principal works to score the lesson, but focuses mainly on choosing refinements and reinforcements that he thinks will be meaningful. He expounded:

I will go to code and look at what that last area of refinement was and what kind of strategies were listed in there. I do that and then of course I look at the rubric and then I will take my evidence and look through after I score it just to make sure that I feel comfortable with areas of refinement and areas of reinforcement. And then for that area of refinement, then I will actually try to go through and try to find some resources.

This commitment to ensuring a quality post conference and to providing additional professional development has resulted in the principal perceiving a shift in the teacher's views of the evaluation process as a whole:

Three years ago when I first started here as an assistant, there was great trepidation and great dread. Oh here they come again, [the principal] and I were in everybody's classroom at least once a month doing observations or drop-ins or whatever. But then the more formalized [observations], folks really dreaded those. One it was fairly new and two, they really didn't get the purpose. They were more focused on, I think, this idea of gotcha. We worked really hard to kind of dispel that. I think now three years into it, from my experience, teachers are more relaxed. They are more understanding of what this process really is. The example I use, I was in a kindergarten classroom, last Friday. I had to do a final observation. In years past when I've been in that same classroom, the teacher, as soon as I walked in the door, the teacher would freak out, just really get stiff,

and you know, she just moved right on. She was transitioning from one lesson to the next and I was there to see the reading part and she just made that transition just as smooth as you please and it was like I wasn't there in a sense, and that's kind of what you want... I have seen growth specifically in some teachers that needed that growth and now we have a much richer discussion. I guess when we talk about standards and objectives or we talk about presenting instructional content, they have a much deeper knowledge. So that's been the good thing.

All of the teachers surveyed (n=7) agreed that the principal utilized the rubric to provide them with frequent feedback on their performance. All but one of the teachers also agreed that the feedback provided to them as a part of the evaluation process was meaningful (M=3, n=7).

School B

Principal Attitude. The principal at School B views the evaluation process as a way to generate self-reflection. The principal stated, "I think the primary role [the evaluation process] has played is self-reflection...[The TEAM model] gave [teachers] the opportunity to have conversations, not only with the observer, but with each other." Colby et al. (2002) deduced that teachers who viewed the evaluation process as a way to improve their instruction were more likely to make instructional improvements. Generating self-reflection is a structure for improving practice. Simoncini, Lasen, and Rocco (2014) summarized the literature on reflective practice and professional dialogue concluding that the combination of reflection and purposeful professional conversation can lead to actions that generate transformative learning. So viewing the evaluation process as a tool for self-reflection is a way of viewing it as a tool for instructional improvement when that reflection is connected to further action. The principal's focus on reflection itself made it unclear whether the principal viewed the reflections as leading to any

next steps. On the teacher survey of implementation characteristics, the majority of teachers surveyed agreed that the principal views the evaluation system as a way to help them improve their instruction (M=3, n=7).

In addition, Kimball and Milanowski (2009) ascertained that whether or not the principal viewed the evaluation system as a useful tool influenced the extent to which the principal ensured the system was valid and reliable. The principal at School B stated that he views the evaluation system as a way for him to be more aware of what is occurring in the classroom. “I think that it has allowed me as an evaluator to be more cognizant of some of the successes I see in the classroom as well as to be more aware of some of the areas of concern.” The principal also shared concerns that the evaluation observations are not always matching the student achievement data that is reported for teachers, “We may do an evaluation component and the teacher score in the medium or the three range. Yet when they get their growth scores back, they have individual growth scores of a 1.” For the principal this was less of a concern about how accurately they were assessing instruction at the building level and deemed more an indicator that the assessments were unfair or invalid. The principal elaborated, “There are ten, and twelve year veterans that are getting the same number of evaluations as those that are just entering the teaching profession [due to individual growth scores of a 1]... You know we’ve got some good teachers in our building... So you know my heart goes out to educators in that aspect.” Despite this feeling that the disconnect is primarily on the side of the assessment, the principal does use the observation data to work on ensuring reliability.

Perceptions of Fairness. Of the teachers surveyed at School B, 7 out of the 9 teachers disagreed that the rubric was an effective evaluation tool (M=1.78, n=9). Although 6 out of the 9 agreed that the evaluations were an accurate reflection of their teaching ability (M=2.44, n=9).

This indicates some mixed feelings about the validity of the evaluation process. The principal also stated, “I think most of them are very accepting of [the evaluation process] and most of them see the validity of it. I think the biggest concern on the TEAM model is not the model itself, you know, I think their concern is on the growth portion of it.” Not seeing a connection between the observation of instruction and the outcomes for students could cause some teachers to be distrustful of the evaluation process as a whole.

Ensuring a reliable and valid evaluation process is dependent upon evaluators working in an ongoing manner on quality assurance (McLaughlin, 1986). The principal shared that the leadership team meets on a consistent basis to ensure inter-rater reliability.

We have meetings on an ongoing basis. Overtly we’re talking about [our scoring]. Many times during the scoring process, prior to the post conference, we’ll get feedback from each other. We’ll share what we’ve observed. These are our initial scores. What’s your take on what I’m sharing with you? See if you are in agreement or perhaps you have a different perception than I do.

This could contribute to the fact that 8 out of 9 teachers agreed that the evaluators were qualified to accurately assess their instruction ($M=3$, $n=9$).

Positive Relationships. Bird and Little (1985) suggest that reciprocity in the relationship between the observer and the teacher is essential to the evaluation system being used as a tool for professional growth. This reciprocity is seen both in the give and take that occurs in the post conference, and in who holds the responsibility for improvement beyond the evaluation. The teachers surveyed all stated that they use the suggestions provided to improve their instruction ($M=3.33$, $n=9$), indicating that they are doing their part in the reciprocity of the relationship.

When asked to describe the relationship between the evaluator and the teacher, the principal stated:

Collaborative. I think most of them would say, perhaps not, but I think most of them would say that when we get together and talk our goal is for everyone to be successful. We want them to be successful; they want to be successful; and we work collaboratively to do that. We're a small school. We know each other. Everybody knows each other for the most part fairly well. So, we want them to be successful. You know...I think collaboration would be the word that I would use.

The majority of the teachers surveyed, 7 out of 9, agreed that the relationship between the evaluators and the teachers was a positive one. However there was a high degree of variance in that answer, with two respondents strongly agreeing and one strongly disagreeing ($M=2.89$, $n=9$).

At School B opportunities for support are limited and finding that support seems to fall generally on the responsibility of the teacher. The principal states, "Teachers for the most part were to themselves and now they have to step out to try and be as successful as possible." Of the teachers surveyed, 5 agreed that the evaluators and the teachers work together to improve instruction and 3 disagreed ($M=2.75$, $n=8$).

This notion of the teachers being responsible for seeking out help is also noted in the principal's description of avenues teachers with low individual growth scores have taken to try to improve:

They seek out for help on different levels, not only at the school level, but their peers, across the district, the core office in Knoxville; they've made contact to try and get some assistance; and still they are scoring at that level.

For teachers seeking help with their personal goals, the principal describes opportunities for observations of other teachers that have been offered. “We’ve made opportunities available for the teachers to view other teachers in the building that they think are successful. Or we’ve also made arrangements for them to visit other schools in the district to be able to watch some of the teachers.” In these cases the responsibility for improvement is still falling solely on the teacher and what they glean from the experience.

The principal also explained that there is the opportunity to utilize PLC time and after school meetings to focus on teacher needs related to the rubric, although it is unclear how often these meetings might be utilized for this purpose. He explained that a few years ago these meetings were used a lot for “healthy discussion” around components of the rubric, but not used as much for that purpose now. However, they do have the “opportunity to get that group of people together.”

High Quality Feedback. Providing high quality feedback begins with the preparation of the post conference plan. Danielson (2012) highlighted the importance of the post conference as an opportunity for a reflective conversation, while Conly and Glasman (2008) suggested the additional need for concrete suggestions over generalized statements about instructional effectiveness. The principal at School B described his process for preparing for the post conference with a focus on scores saying, “I pull down their TVAAS scores and their growth scores and their achievement scores. I want to have everything that I can have accessible so when I bring out an area of concern or an area of success, I’ve got data to be able to back it up.” While the principal does reference the post conference format of providing a reinforcement and refinement, “We talk about the successful things that we saw and that we saw areas of concern.” He focused heavily on the scoring aspect of the post conference in his description adding:

We don't share individual scores out at that time. We do give them an average. We'll take the 12 aspects of the [instructional rubric] and we'll average that out to show them what the overall average was for the observer and we'll share what their self-score average was...If it is a large discrepancy, then we'll have a discussion about the large discrepancy as a whole, but if [the scores are] pretty well in line with each other, then we'll tell them that as a whole, overall we are pretty much in agreement. There may be some portions where...there may be discrepancy, but overall that's good.

Despite the lack of focus in the principal's answer on the quality of the feedback provided in the post conference, teachers reported that they found the feedback they received during their evaluation to be meaningful to them ($M=3.11$, $n=9$). The majority of the teachers also reported that they were provided with specific suggestions for improvement ($M=3.33$, $n=9$).

School C

Principal Attitude. The principal at School C views his role in the building as an instructional leader and the evaluation process as an essential component of the school culture. The principal contends that the evaluation process at his school is "all encompassing," stating, "We use [the evaluation rubrics] to drive data, to drive our decisions, to look for good teaching, and it's used as a whole approach to our school week and our school day." He further described the connection of the evaluation process to other facets of daily life in the school saying, "We use [the evaluation data] on a daily basis to [evaluate] what we've done in cluster (the school's job-embedded professional development model) and our TLTs (TAP Leadership Team meetings)...[to] look at our reports and see where our scores are at, where our refinements and reinforcements are and try to address them in TLTs, through cluster, and then out into the classroom." The principal at School C's reported view of the system as an integrated facet of

daily life within the school suggests that he views the evaluation process as a valuable tool for instructional improvement. Kimball and Milanowski (2009) contended that principals who viewed the evaluation system as a valuable tool would also spend time ensuring that the system was valid and reliable.

The leadership team meets regularly, once a week for an hour, and they use the evaluation process as a springboard for supporting teachers in the building. The principal described this level of differentiation specificity saying,

We may pull just the refinement piece...and look at it and say ok, we're missing something with assessment, for example, and so we...try to break it down into the cluster groups...Where do we need to go with that? Is it because they're not making something that is measurable? Are they leaving it completely out? Do they assess during the instruction but they didn't plan for it? Or vice versa, did they plan for it but then didn't do it during the instruction piece? [Then in cluster] we'll get the rubric out...and we'll say here's what the rubric says. Here's what from our walk-throughs, from our drop-ins, through the actual evaluation piece, what are we seeing, what is happening.

The principal clearly articulates a connection between the evaluation processes, the professional development provided to teachers weekly, and the walk-throughs and drop-in visits he conducts as an administrator. The teachers surveyed agreed strongly that the principal views the evaluation process as a tool to help them improve their instruction ($M=3.75$, $n=4$).

Perceptions of Fairness. The teachers surveyed at School C agreed that the TEAM/TAP rubric was an effective evaluation tool ($M=3.67$, $n=3$). They also agreed that the evaluations were an accurate assessment of their teaching ($M=3.67$, $n=4$). The principal also stated that he felt teachers in the building believed the evaluation process to be fair.

I think now [teachers] understand that it is evaluating teaching. It's not evaluating them as a person... I have had very few complaints about evaluations in my building... I think that they think that [the evaluation process] is very valid. I think that they think, that they know that we are looking for teaching and not themselves and that's taken the personal element out of it. I think that if that evaluation process has that personal feel, then you don't get a true feeling of validity. It may be valid, but you may get resentment that it's about them personally instead of the teaching itself.

Weber (1987) discussed the importance of teachers having confidence in the evaluation system in their willingness to utilize the feedback from the process. The principal at School C affirmed his teachers' trust in the process stating, "I think the confidence of our teachers knowing that it is reliable because the team has worked on reliability and with us taking back to them that in cluster, that our inter-rater reliability is .68 apart... So that's an indicator of the confidence that everyone is looking for, everyone hopefully is looking for the same things. Looking at the same criteria." The teachers surveyed indicated they indeed have this level of confidence with all of them agreeing that the evaluators are qualified to accurately assess their instruction ($M=3.5$, $n=4$).

In order to generate inter-rater reliability the team works consistently to ensure a common understanding and to communicate the work they have done to teachers in the building. The principal explained that the team has worked on inter-rater reliability or issues related to the evaluation process in half of the TLT meetings they have conducted this year. The team works to ensure that they are going to a depth of learning in TLT that will change something else in the building as a result of their meeting. The principal describes that work saying:

We in TLT, we bring that up [in CODE] and talk about it...We will pull a national lesson off every so often and we'll score it and let them tell us what the national raters rated that lesson to make sure we're all still seeing the same things...[Sometimes] We break down and break it into clusters of where the refinements and reinforcements are showing up...Are there things that are the same, are there things that are totally different? Is it something because its grade level specific or is it whole school...so, trying to get deep with it. Just pick that thing that we know we are having a descriptor or an indicator that is struggling a little bit, lets get deep with it and see what we can find out about it.

Not only is the team working in an ongoing manner, with a regularly scheduled weekly meeting time, but within that work, they are working to push to depth of understanding with small pieces of the evaluation rubric at a time. Sometimes that work is through the lens of better scoring and sometimes it is through the lens of ways to better support teachers in their practice.

Positive Relationships. A high level of responsibility for the support and growth of teachers is held by the principal and developed in the school's leadership team meetings. This support includes differentiation of the professional development in cluster, as described above, but also in the individual follow-up that is provided for teachers following their evaluations. Weber (1987) described effective evaluation systems as ones that functioned from a shared culture of collaboration and improvement. The ownership of School C's leadership team on the improvement of teacher's demonstrates a high level of commitment to their end of this reciprocal relationship. The principal described this relationship saying, "[The teachers] know that they are going to have one reinforcement and one refinement, and we're going to work on that, we're going to follow up with that, and hopefully in each cycle that the reinforcement or refinement, hopefully those are not the same things that show up on a person from evaluation to evaluation."

The teachers surveyed also report that the teachers and the evaluators work together to improve instruction (M=3.5, n=4) and all four teachers strongly agreed that the relationship between the evaluators and the teachers is positive (M=4, n=4). The principal also shared his own commitment to ensuring high quality support and follow-up from the evaluation process:

I spoke with [our NIET support person] about myself personally. [I asked,] How can I follow-up with everyone that I'm evaluating and how can I follow-up at the level that I want to follow-up with them? She and I had a conversation about [that]. [She said,] lets pick one for sure and follow-up with [that teacher]...don't do five at a very low level, do one at a very high level, and then catch the other four at a level. [Then] hopefully...when they go to the next cycle, they won't be at that lower level for whoever has them the next evaluation.

The teachers also strongly agreed that they use the feedback that is provided to them to improve their instruction (M=4, n=4). So not only is the school's leadership team taking a large share in the responsibility for teacher growth and improvement, but the teachers are taking their share in that responsibility as well. This shared responsibility for growth is evident in the principal's description of his teachers.

I am confident that the teachers here at this school are going to do what they are supposed to be doing, whether it's in an evaluation or if it's not an evaluation and I just walk in for a drop in, or for a follow-up or just to go in just to drop in...I'm going to see the things that we've worked on through our TLTs and through our clusters.

High Quality Feedback. Due to the structures of cluster and follow-up being connected to the evaluation process in School C, more opportunities for feedback connected to the evaluation rubric than in just the post conference itself are provided to teachers in an ongoing

fashion. Still within the evaluation process itself, the vehicle for feedback remains the post conference. The principal described the post conference process at School C as the beginning of a feedback loop for teachers saying, “It’s just a continuous cycle of the follow-up with every person.” The principal stated that one of the areas his team focuses on is ensuring that they capture what students are doing so that they can incorporate that student evidence into their feedback. He stated:

That’s one of the things that we really try to focus on is scripting the students, not scripting the teacher. What are the students doing and how are they reacting? Do they know their roles and do they know what they are doing?

In planning for the post conference, the principal included a focus on ensuring that he provided a suggestion saying, “With the refinement. I will use the terminology of like, this is something we can refine upon...I’m not saying that it was bad, but this may be something that we need to look at, and go forward and give suggestions.” The principal connected those suggestions to the creation of a follow-up and support plan, indicating that the post conference was not the end of the evaluation process, but the beginning of a process of support for teachers. The principal added, “Then [sharing] the follow-up process. Then knowing that we follow-up, we have the post conference; we’ve talked about a refinement. Then they know that they’ve got some support there.”

The teachers surveyed agreed that their evaluators provided them with specific suggestions for improvement ($M=3.5$, $n=4$) and that the feedback that was provided to them was meaningful ($M=3.5$, $n=4$). In addition, 3 out of 4 teachers agreed that the principal utilized the rubric to provide them with frequent feedback on their performance.

Implementation Summary

The characteristics of implementation are unique to each school. The attitude of the principal, the perceptions of fairness, the relationships between the evaluators and teachers, and the quality of the feedback that are provided create the culture of evaluation that is present in each school building. The quantitative analysis that will follow will show the potential impact these differences might have on teacher efficacy. In addition, the description of these characteristics will provide a framework for attaching meaning to the qualitative analysis of the teacher's interviews.

Research Question 1: To what extent does implementation of TEAM act as an intervening variable for teacher self-efficacy? (QUAN)

The previous description of the implementation characteristics at each school provides a frame of reference for the quantitative data analysis. Due to the limited number of respondents, statistical analysis could not be completed; however, the following analysis will still provide some explanation of the potential for differences in TEAM implementation to act as an intervening variable for teacher efficacy. Additional research would be needed to verify any trends that may be present in this limited data pool.

There are differences in the mean scores for the three schools on both the pre-evaluation assessment and the post-evaluation assessment. On the pre-assessment, teachers at School A had a mean self-efficacy score of 3.07 (n=7). Teachers at school B had a mean self-efficacy score of 2.74 (n=9), and teachers at School C had a mean self-efficacy score of 3.48 (n=4). On the post assessment, teachers at School A had a mean self-efficacy score of 3.39 (n=5), teachers at School B had a mean self-efficacy of 2.67 (n=4), and teachers at School C had a mean self-efficacy of 3.72 (n=2). As described in the characteristics of implementation for each school, School C had

the most cohesive culture of evaluation, with structures for professional development and support tied to the evaluation process. This suggests that there could be a correlation between evaluation implementation practices and teacher self-efficacy overall.

For teachers who completed both the pre-assessment and post assessment, there was some change in the mean averages. School A had a mean change of 0.23 (n=5), School B had a mean change of 0.09 (n=4), and School C had a mean change of 0.12 (n=2). Some change in efficacy did occur for all of the teachers in the sample. The qualitative data analysis in the next section will provide some insight into how differences in the characteristics of implementation may have influenced these changes.

For all of the teachers who completed the pre and post assessment, the average change in efficacy was 0.16. However, an item analysis was conducted to see how many items changed from the pre to the post assessment. Interestingly, all of the teachers changed beliefs on several items. Most of the changes were single point changes. These included both positive and negative changes.

One of the teachers interviewed described her changing beliefs saying, “Today I feel better...[That day] we took the SAT 10 and I just felt depressed about it...but today if I would have taken it I would have felt a little better.” Another teacher in this study described the changes as continual strengthening. She stated, “I don’t feel like [the evaluators] are coming in and telling me to fix anything... They are coming in to help me strengthen something that’s already in place but could be a lot stronger.” So, while the teachers overall efficacy may not have fluctuated very much numerically, the makeup of those beliefs changed (see Table 8).

Table 8.

Teacher Self-Efficacy Item Analysis

Teacher	Number of Items Changed	Number of Positive Change	Number of Negative Change	Number of Items Changed More than 1 point	Self- Efficacy Change
Teacher 1	18	17	1	3	0.63
Teacher 2	10	9	1	0	0.27
Teacher 3	6	3	3	0	0
Teacher 4	8	3	5	0	-0.07
Teacher 5	14	11	3	1	0.3
Teacher 6	8	7	1	0	0.2
Teacher 7	21	12	9	3	0.13
Teacher 8	10	2	8	1	-0.24
Teacher 9	10	9	1	0	0.26
Teacher 10	6	4	2	0	0.07
Teacher 11	9	7	2	0	0.17

Research Question 2: How do differences in the implementation of TEAM influence teacher self-efficacy? (QUAL)

The previous two sections demonstrate that there are differences in the implementation characteristics at the three different school sites and that teacher's self-efficacy beliefs did fluctuate over the course of the semester. This section will answer the question: How do differences in the implementation of TEAM influence teacher self-efficacy? Three teachers were interviewed and their answers will provide some explanation as to how different implementation characteristics may influence teacher's self-efficacy beliefs. These findings are limited in their generalizability, as they represent the experience of only these three teachers.

Differences in the implementation characteristics of TEAM have varying influences on teacher efficacy. Three major differences seemed to have the largest impact on efficacy for these three teachers: when feedback was connected to students, the amount of stress generated or not generated by varying levels of support beyond the evaluation, and the teachers' perceptions of the fairness of the evaluations, as well as their perceptions of the principal's attitude about the process. The following section will explain each of these differences and provide examples from the teachers of how these differences influenced their beliefs.

Feedback Connected to Students Generates Efficacy

All three of the teachers in this study felt the most beneficial part of the evaluation process was being able to receive feedback on their instruction. However, they also all described instances where they had both utilized the feedback they had been provided and where they had dismissed it.

When the teachers described examples of feedback that had resulted in behavioral changes, the feedback was connected to student data and/ or the evaluators had tried the

suggestions in their own classrooms. These connections to student outcomes could be creating efficacy by generating vicarious experiences for the teachers. Bandura (1977) described a vicarious experience as seeing others succeed at a task. The addition of references to the impact on students when providing suggestions or feedback may push the feedback to become a vicarious experience. One teacher described the feedback she was provided and the impact it had on her instruction saying:

She was able to show me examples of how it helped [her] students. You know, the assessment piece really helped them...She explained to me what assessment was, how assessment should be in my lesson plan, how I needed to plan for assessment, what I needed to do, and then after she did all that, I began to use what she talked to me about.

So then the next time I had an evaluation, I did that, and that became my reinforcement.

Bandura also states that the more similar the model to the observer, the more impactful. This resonates in some of the teachers' comments that the evaluators with experiences most similar to theirs had the most impactful feedback. The ability of the evaluator to connect his or her own experience in the classroom to the teacher's experience appears to connect that feedback back to students generating a vicarious experience for the teacher.

In contrast, when the feedback was disconnected from students, the teachers had difficulty trusting the feedback. The teachers were uncertain about the feedback when they felt the evaluators didn't know their students or understand their grade levels. In these cases, the feedback failed to generate a change in the teachers' beliefs about that area of instruction and failed to foster a change in their behaviors. For example, one teacher said:

I sometimes do not feel like the evaluator is familiar with this age group. So when they say [students] can do such and such, I disagree because I know these kids better than they do and I feel like they're not familiar with what's age appropriate.

However, for one teacher, evaluators were able to overcome this disconnect by focusing on connecting the feedback to students and to research that is age appropriate:

Both of the [peer evaluators] that I had this year were high school [teachers]... When they came in, the way they talked about their students and their care for their students. Then the data that they brought to show me, and the information, it was on my level. I mean they had researched themselves to get it where I could actually use it! Here it is. This is what you could do... I don't feel like I have to do it, but I feel like because the data that they've brought in and the research that they brought in, when you look at it, you're like, wow! That might work. I might not want to do that one. That's not for me, but I could do this, this, and this.

Furthermore, when little or no feedback was provided, there was also no impact for the teacher.

One teacher described a previous school experience with the TEAM evaluation saying:

There was no feedback before. They just came in and said, 'This is what I chose. This is why.' They didn't bring any research; they didn't bring anything to tell you, 'Well, you could try this. What about trying this?'... I didn't do anything with the other. I'll just be honest. You were called in the office. 'Here's what it is.' I'm like, 'Okay.' Take it. Filed it. I didn't do anything with it. We weren't required to do anything with it.

Based on the experiences of these three teachers, feedback provided during the post conference may have the potential to generate a vicarious experience when it is connected to students through either research or the evaluator's own experience. When these sources of

efficacy are generated, the teachers tend to try new things in their teaching practice and in trying them, sometimes generate mastery experiences. This combination of efficacy sources leads to changes in their efficacy beliefs. In contrast, when the feedback is not connected to students, the feedback may have little or no impact at all.

Support Beyond the Evaluation Process Influences Stress Response

The teachers stress level as a result of the evaluation process may also generate a source of efficacy that can lead to increases or decreases in efficacy. The level of support provided to teachers following the evaluation process seems to have influenced the amount of stress two of the teachers felt as a result of the evaluation. While the third teacher stated that the evaluation process itself didn't cause her any stress saying, "I felt comfortable with [the evaluations] because I like for [the evaluators] to come in and see what we're doing. You know, so it doesn't bother me, but I've heard from a lot of people that it stresses them out."

The teacher's stress response to the evaluation process may generate an affective state that can either promote efficacy or serve to lessen it over time. Bandura (1977) theorized that the way a person copes with and handles stress may have a significant impact on their motivation to complete a task and that performance expectations can be negatively impacted when stress causes a negative emotional response.

For two of the teachers the opportunities for support seemed to be connected to the amount of stress they expressed during the interview process. For one teacher, a lack of support led to a continuous stream of non-mastery experiences that left her frustrated when she couldn't find a way to fix the concerns highlighted in her evaluation process:

It is just trial and error...[The evaluators] are saying this is wrong so, you know, I have talked to other peers. How would you do this? What would you do? Did you score well in

this? If they did, then they would offer suggestions. But, you know, sometimes they didn't have any suggestions...I think I've gotten better just from some of [the evaluator's] suggestions. On some things I don't feel [like I've gotten better]. I feel like I get marked down every time. Sometimes I feel like I'm not [getting better]. I mean there have been times, I mean, if they say this is a refinement, then I work. I don't want that to show up.

The teacher's desire for improvements, coupled with limited opportunities for support contributed to the stress that was generated as a part of the evaluation process. This negative stress may have contributed to a decrease in the teacher's efficacy beliefs.

For another teacher, an abundance of support led to her feeling strengthened and accomplished. The presence of a support structure seemed to limit the amount of stress that developed as a result of the evaluation process and resulted in her having little to no negative stress response from the process. She described the support saying:

I don't feel like they are coming in and telling me to fix anything. I don't feel like that's what they are trying to do. They are coming in to help me strengthen something that's already in place but could be a lot stronger...After the actual evaluation, they don't just leave you...I like that cause it's more like you are supported.

The teacher described the support as coaching, weekly professional development where she learned strategies, collaboration with peers, and conferences she was able to attend. All of this support led to her feeling that she had improved over the course of the school year, "I am very excited because I feel like I've grown more this year than I have in a long time, because of those things that are in place."

All of this support created a positive environment that limited the amount of stress this teacher felt as a result of the evaluation process. This positive affective state seemed to keep her efficacy overall at a generally high level, with some areas improving over the course of the semester. These supports may also have provided other sources of efficacy such as vicarious experiences through the modeling in cluster or suggestions given from peers, verbal persuasion in the coaching, or the development of mastery experiences when she tried the new learning in her own classroom.

For the other two teachers, support was something that they desired but didn't know how to attain. They both discussed the support they would like to see provided.

I wish there were more conferences they could give us to go to...I feel like we don't get [enough support.]...I would really like to talk to another [primary teacher.] Not just any teacher- I want somebody that has received great scores. I mean not great scores on her evaluation, I'm talking about her test scores are great and her students are phenomenal. I want to talk to somebody like that. Or send me to a conference to help me get better...I'd like to have help and I don't feel like we have the resources.

The other teacher reiterated this saying,

I don't think [teachers] get enough of going to inservices, seeing these new ideas and bringing it back. I think we need more professional development...I would like to go to even the schools in the county to see how they are teaching this. You know, just more one on one. Let me go to [other schools] and let me see what they are doing with grouping, or reading, or math. Give me an idea. I would like more of that to happen.

All of the teachers noted that more support was necessary to make the evaluation process more beneficial for teachers. Structures for support generated mechanisms that assisted teachers in

managing stress. Without these structures negative stress responses became capable of generating negative impacts on teacher efficacy.

Perceptions of Fairness and Attitude about Evaluation Give Power to the Sources of Efficacy

While the feedback and support provided during the evaluation process generated sources of efficacy, the teacher's attitude toward the desire to grow and change was also instrumental in how much stock the teachers put into the evaluation feedback they were given. All of these teachers talked about being open and willing to accept feedback. One of the teachers said, "If you're willing to look at things; look at your scores, look at your shortcomings, and try to fix them...I think if you are willing to work on it, and you are really willing to study what you are doing, that you can become a great teacher."

Each of these teachers had this same attitude about their instruction and desire for improvement; however, there were times that they did not use the feedback they were given. In the previous section it was noted that when the feedback was disconnected from students, the teachers tended to not utilize the suggestions provided. The teachers' perceptions of fairness and their attitude around evaluation either provided the evaluation process with the power to generate sources of efficacy, or it stripped that power away, making the evaluation process have no impact at all. Two of the teachers described times that they disagreed with the feedback they were given or disagreed with the scoring of the lesson and that seemed to influence their willingness to give the evaluation process the power to generate sources of efficacy for them. In one example, both teachers shared cases where the context of what was observed in the lesson caused them to feel their evaluation was inaccurate. They both described a "bad day" and stated that they didn't think the evaluation should have been conducted.

I mean, you have a child that throws up in the room, and I've had that, or, uses the bathroom on himself, and you've got interruptions, and sometimes you just have to say, 'Look, we're just going to have to do this another day.'

The other teacher had a similar experience.

I did not agree with what he had because it was a horrible day. I had a child with a nosebleed; I had to help three other children learn how to blow their noses, because they were having major allergies. I have a little girl in here that has special needs with behavior. She went off. I'll be honest, if I had been in the classroom, I would've said, I'll be back. I'll come back another day.

The teacher went on to explain that because she didn't think the feedback was valid she did not use it to make any changes to her instruction. She said, "To me, that wasn't a weakness. That's not something where I have a problem. Maybe if somebody else comes up and says that, then I'll look at it even more..."

One of the teachers also explained what happened when the attitude about the evaluation process went from being just something that had to be completed, to the process being about teacher growth. A shift in this attitude caused the evaluation process to go from having no power to generate sources of efficacy to having a large impact on her as a teacher. She explained:

Before, [the evaluators] didn't ask you to change [your evaluation]. They didn't ask you to work on it. It was just something they had to do. Here it is. I filed mine. You file yours. Nobody ever came back and said how's it going? In the other model, I didn't see a rubric. I didn't even know what I was supposed to be doing. I was just doing what I was doing. There wasn't any [support]... I've [taught] for 22 years and this year I am very excited because I feel like I've grown more this year than I have in a long time, because of those

things that are in place...At this school, they explain the rubric. They explain to us why things are important. They made connections for us.

Another teacher described how viewing the rubric as a set of expectations and a target for perfection made her feel stressed and overwhelmed. When the attitude of the evaluation process was that the rubric was a measure for perfection rather than a tool to strengthen her practice, this attitude generated a negative source of efficacy, despite the teacher's willingness and desire to improve her craft.

I mean there have been times if they say this is a refinement, then I work, I don't want that to show up, but I get this fixed and then oh, I've got to find something else. So it's not like I have a perfect score at the end. You know what I'm saying? I'm still working on things...Well, sometimes, you know, I feel like on the rubric there's so many things that you have to do to even get a great score...You know. It's like I'm stressed because I'm trying to hit everything on that rubric. Is that a reflection of what happens everyday? Every lesson? Do I do that every lesson? No. Do I want to? Yes. But I feel like there's not enough time in the day to do all of that. So, I mean, I just feel like it's impossible. Could you get a three? Yes, but if you are shooting for a five, no. You can't do it all.

The perceptions of fairness and the attitude around evaluation did not appear to generate mastery experiences, vicarious experiences, or verbal persuasion on their own, while they did contribute to the teachers' affective states. However, the teachers' perceptions of fairness did appear to influence their willingness to grant the other implementation characteristics of feedback and support the power to generate the additional sources of efficacy. When those perceptions of fairness were not present, the teachers did not pay any attention to the feedback they were provided.

The attitude around evaluation also seemed to influence the power the other sources of efficacy were given by establishing the culture for the evaluation process. The attitude seems to have a connection to the stress the teachers feel about the evaluation process from no stress as a result of no emphasis being placed on the evaluation, to positive feelings of support and growth when the process is seen as a growth tool, to high levels of stress when the attitude around the evaluation process is to strive for perfection.

Conclusion

The three schools sampled represented three varying implementations of the TEAM evaluation model. Each school utilized the tools provided with the model in varying ways. The quantitative analysis indicated that there were differences in the mean efficacy of the teachers surveyed for each school. However, limited data made it difficult to draw any conclusions as to whether these differences were significant.

Analysis of the teacher surveys did provide some insight into how the varying implementation characteristics influenced the teachers' self-efficacy. Feedback that was connected to students impacted the teachers' beliefs and their actions, whereas feedback that was disconnected from students generally had little impact on either. Support centered on the evaluation process influenced the amount of stress teachers felt. Those stress responses influenced the teachers' self-efficacy beliefs. The teacher who felt supported had little stress response and her efficacy remained high, whereas the teacher with little support had a high stress response and her self-efficacy lowered over the course of the study. Finally, perceptions of fairness and the attitude of the principal provide power to the sources of efficacy. In order for the teachers to allow the sources of efficacy influence over their beliefs, they had to believe the

evaluation was fair and that the principal had expectations that they would use the evaluations as a tool for instructional improvement.

While these findings represent the experience of three teachers with the TEAM evaluation system, more research is needed to confirm whether these findings are true for all teachers. Despite the limited amount of data discussed here, several suggestions can be made for improving the process in schools across the state. These suggestions will be discussed in the next chapter.

Chapter 5

Discussion and Implications

The Tennessee Educator Acceleration Model is an evaluation system intended to serve as both a summative accountability measure and a formative tool to improve teacher effectiveness. However, teachers across the state have varying perceptions about the influence that the system has on their teaching practice. Bandura's (1977) theory on efficacy suggests that to change behavior, efficacy must first be strengthened. So, in order to change teacher behaviors, TEAM might first need to improve teacher self-efficacy, the beliefs that mediate teacher behavior. This study attempted to explain the extent to which differences in implementation of TEAM acted as an intervening variable for teacher self-efficacy. In addition, the study examined how the evaluation implementation characteristics influenced teacher self-efficacy by studying the characteristics of implementation through the lens of Bandura's four sources of efficacy: mastery experience, vicarious experience, verbal persuasion, and affective states.

Discussion of Findings in Light of Research on Efficacy

For the majority of the teachers surveyed self-efficacy was relatively stable or increased slightly over the course of the semester. Wheatley (2002) suggested that teachers needed to first doubt their efficacy before they could improve it in order to see a change in practice. However, the relative overall stability of the teachers' efficacy in this study, particularly the interview participant whose efficacy remained stable, yet explained that her practice had improved greatly over the course of the year, seems to suggest that teachers don't necessarily need to doubt their ability to enact certain behaviors to change them. The maintenance of the teacher's high efficacy beliefs resulted in positive perceptions of the experience overall.

Another teacher in this study experienced an overall decline in her self-efficacy over the course of the semester. In Tschannen-Moran and McMaster's (2009) study on the impact of professional development on teacher efficacy, conclusions were drawn that a lack of support for the initial professional development provided in that study resulted in a decrease in teacher efficacy. The same conclusion can be drawn from the findings in this study. A lack of support for the initial evaluation feedback resulted in a decrease in the teacher's self-efficacy. In this case, the lack of support resulted in the generation of a negative stress response or an affective state that served as a detractor to the teacher's self-efficacy. The teacher's individual attempts to change practice were met with continual non-mastery experiences that resulted in feelings of frustration and anxiety.

Teacher efficacy development, therefore, may be enhanced by an attempt to provide stability to teacher beliefs over time. Changing teacher practices may be possible by ensuring a culture of support that enhances and stabilizes teacher efficacy.

The majority of teachers in this study remained relatively constant in their overall self-efficacy, however, the make-up of their efficacy beliefs shifted slightly over the course of the semester. The teachers' scores on different items changed while their overall scores only varied slightly. In their work to develop a conceptual framework for teacher efficacy, Ashton, Webb, and Doda (1983) concluded "teachers' sense of efficacy is negotiated daily in their myriad transactions with students, peers, and administrators" (p. 38). However, Bandura (1997) suggested that efficacy is resistant to change once it is established. The teacher survey results and the interview response suggest that teacher self-efficacy, the teachers' belief in their ability to effectively implement specific teaching behaviors, may be negotiated on a regular, perhaps daily, basis. While this may not be noticeable in the overall efficacy average, the frequency of change

can be seen in the shifting beliefs related to particular behaviors on the two different survey administration dates.

The teachers' self-efficacy, overall belief of their ability to enact specific teaching behaviors may also be driven by their overall teacher efficacy beliefs. Dellinger, Bobbett, Olivier, and Ellett (2008) defined teacher self-efficacy as a teacher's ability to complete specific teaching tasks and teacher efficacy as a teacher's belief in their ability to affect student performance. The teachers' beliefs about these different teaching tasks may be influenced by their beliefs about their ability to influence student achievement. All of the teachers interviewed found the connections to student data and to student outcomes as necessary for them to accept and utilize the feedback from the evaluation process. All of the teachers interviewed in this study also held the belief that they could improve student outcomes by improving their own practice. The need for connections to student data for these teachers who appear to have high General Teacher Efficacy (GTE) could suggest that a teacher's GTE or the teacher's general belief in the ability of a teacher to influence student achievement, as defined by Gibson and Dembo (1984), may be an underlying characteristic that is necessary for a teacher to focus on making instructional changes. The connections between these two beliefs could provide another influencing factor in the amount of influence various efficacy sources have on a teacher's efficacy development.

Finally, generating the sources of efficacy alone does not seem to be sufficient to generate efficacy development. Teachers must also allow these sources of efficacy the power to influence their beliefs. Within the evaluation process, power to generate efficacy was removed when the teachers did not perceive the evaluations to be fair, when they didn't view the evaluation process as important, or when the feedback wasn't connected to students. The

importance of the teacher's ability to grant the sources of efficacy the power to influence their beliefs and thus influence their actions, may be important for understanding the components that are essential to generating change within schools. Previous studies of the effects of various sources of efficacy have explored the impact of the application of the source of efficacy on participants. For example, in a study of male and female students at North Texas State University, Weinberg, Gould, and Jackson (1979) focused on how the use of verbal persuasion impacted the students' performance on a muscular endurance task. Similarly, in teacher efficacy research, many studies focus on the impact of various efficacy sources on participants or on how various education efforts, such as professional development or teacher collaboration generate the sources of efficacy (Chong & Kong, 2012; Tschannen-Moran & McMaster, 2009; Tschannen-Moran & Woolfolk Hoy, 2007;). These studies fail to take into account the teacher's role in receiving the source of efficacy, generally considering the application of the source of efficacy as a treatment that is consistent across all teachers. The teacher's role in receiving that source of efficacy is a new variable that warrants consideration.

A Support Centric Model for Teacher Evaluation

Principals interested in improving instructional practice should consider the role the evaluation process plays in affecting teacher efficacy and behavioral change. To generate a culture of continuous growth and improvement for teachers, the evaluation process may need to be part of a cohesive teacher support structure that is an integrated part of the daily functions of the school day. By generating a support centric evaluation model, principals can improve the implementation characteristics of evaluation and ensure that teacher efficacy is maintained and strengthened as they support positive instructional changes.

Currently, schools seem to fall along a continuum of experiences when it comes to their evaluation practices. Schools at one end of the continuum are task-oriented, where teachers perceive the process as “just something we have to complete.” While teachers in schools at the other end of the continuum view the evaluation process as an integrated function of school life, where the language and the process is “just a part of how we do school.” In shifting through these phases, some schools can begin to focus on communicating the use of the evaluation process as a growth tool, but not prepare for ways to support teachers in making instructional changes. When this happens, teachers are left to fend for themselves to find ways to improve. As noted in Chapter 4, this can cause teachers stress that results in decreased efficacy. However, when just seen as a task to complete, the evaluation process also can have limited or no power to influence teacher behavior. Therefore, shifting from task-oriented to growth-focused is essential for changing the usefulness of the system if it is to be used as a catalyst for changing teacher practice. To make that shift without negatively impacting teachers, principals should consider a model for evaluation practice that focuses on support as the central element to their evaluation practice design, rather than accountability.

The current expectations around teacher evaluation are heavily tied to accountability. Shifting the purpose for evaluation can result in changes to the implementation characteristics associated with an effective evaluation process. In addition, shifting the focus to teacher support can also ensure that teacher efficacy is not negatively impacted over time. The focus of the model can impact the implementation characteristics by providing focus to how teachers and administrators interact with and use the evaluation system. These connections can be seen in the Figure 4.

When the evaluation process is accountability centric, there is little drive to connect the evaluation to any additional structures. In this model, the role of the evaluation is simply to assess the quality of the instruction. Any work the principal does in response to the data collected is to ensure accuracy of the data for accountability purposes. The primary focus of the evaluation is the evaluation itself, and so any growth that is expected to occur from the process becomes left to the teacher to construct. The feedback provided is generally limited to the evaluation itself, and sometimes focuses heavily on scores. When the teachers only experience the evaluation language during an evaluation, this limited exposure can cause distrust of the process and generate perceptions of unfairness amongst teachers. Depending on the extent to which value and expectation is placed on the amount of work teachers are expected to do independently within the evaluation framework, this model can also cause teachers stress and anxiety.

In a support centric evaluation model, everything becomes focused on ensuring that the process results in supporting and helping teachers as a result of and throughout the evaluation process. Changing to a support centric model for evaluation will require principals to reassess how they are using the evaluation process within their buildings and school day. This will also have positive impacts on the implementation characteristics for evaluation.

There are several steps that a principal can take to begin to become support centric in his/her evaluation practices. Each of these steps will have varying influences on a variety of implementation characteristics. First, the principal should consider the structures and people who

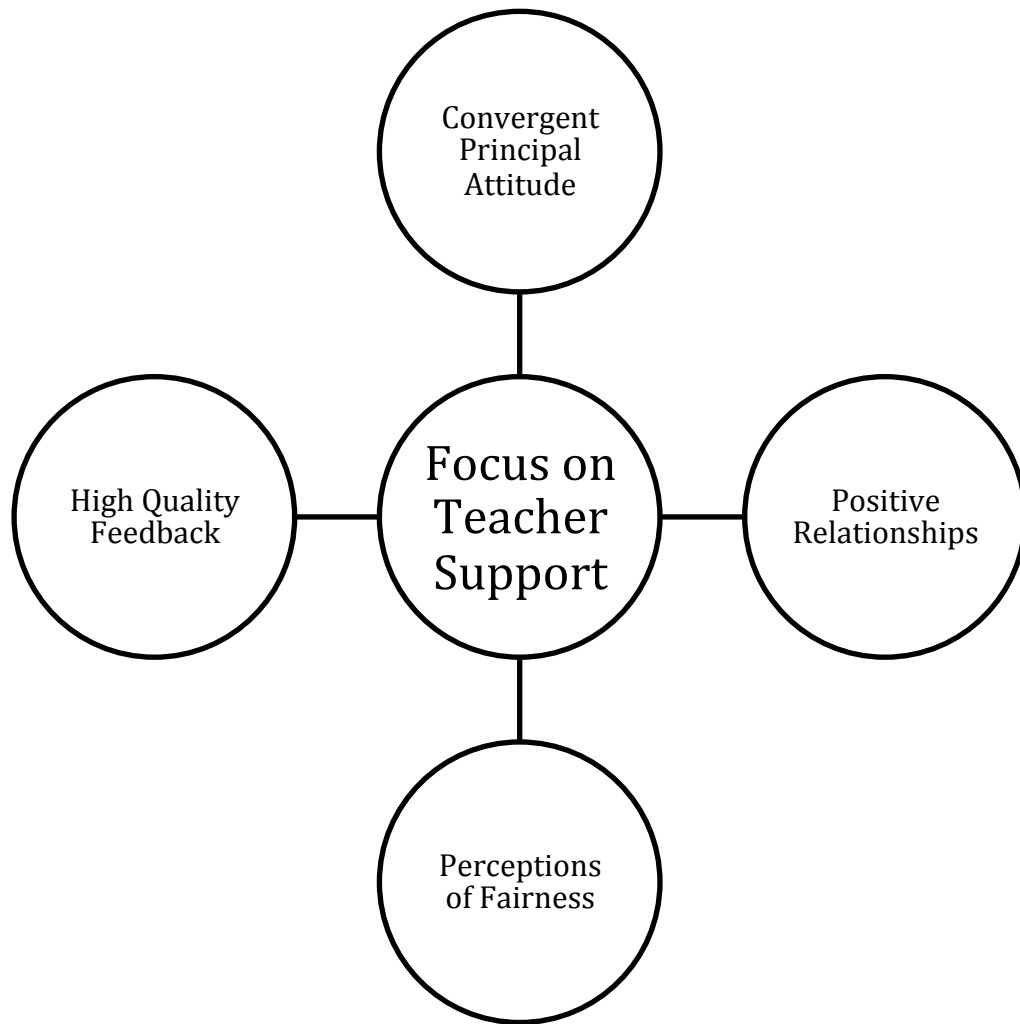


Figure 3. Support Centric Model

can be used to provide formal support to the evaluation process. Developing formalized support roles and formalized structures where support will happen is a way to ensure that support does occur. In this study, the school that had these more formalized roles and structures was able to generate a support centric evaluation model. The establishment of both the roles and structures added a degree of accountability ensuring support for the evaluation process would occur on a regular basis. Including evaluators within these formalized roles for support forms a mutual relationship between the evaluators and teachers, where both are responsible for the teacher's development and growth through the evaluation process. Expanding the use of peers as evaluators or including administrators in the delivery of professional development could also establish more opportunities for shared responsibility of teacher growth. Combining both evaluators and teacher support personnel into one instructional support team might be one way to develop a cohesive evaluation support team. Connecting evaluators to support structures has a direct impact on the relationships of the teachers and the evaluators, developing more reciprocal relationships related to growth from the evaluation process. The formalization of the structures and roles for support also increases the consistency of the feedback provided. The frequency of support provided through these support structures and formalized support roles may increase the perceptions of fairness as teachers are exposed to the language of the evaluation on a more regular basis and their understanding of that language is developed over time.

Principals can further develop evaluation practices that are support centric by ensuring that there are plans for the instructional support team to meet regularly to examine data from the evaluation process. In these meetings, the instructional support team may work on inter-rater reliability to confirm that teachers are receiving consistent and accurate feedback that supports

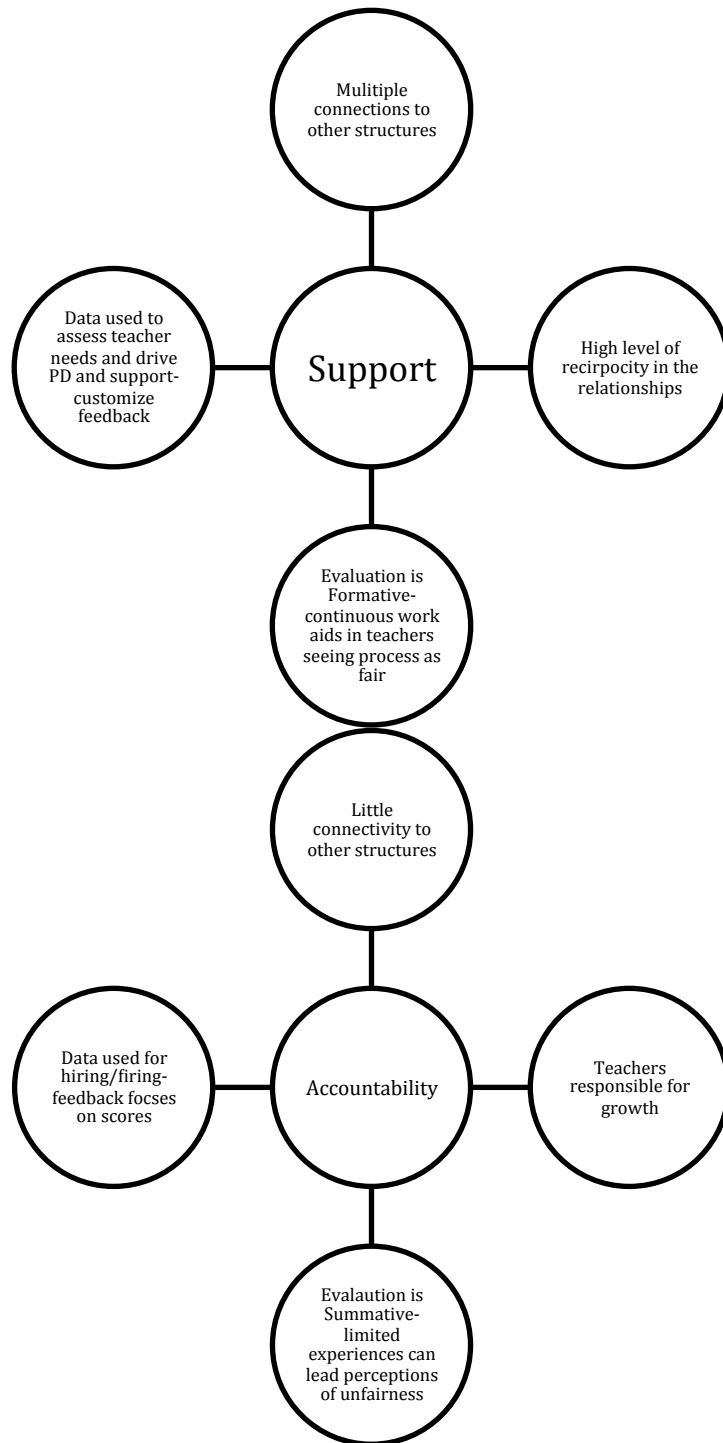


Figure 4. Support Centric Model Versus Accountability Centric Model

their practice. They may also use that data to determine the needs that will be addressed in the formal professional development structures they have designed to add support to the evaluation process. Utilizing the instructional support team and the data from the evaluations as a decision making team for the focus of the support guarantees that the evaluators take active responsibility for ensuring teacher growth as a result of the evaluation. Monitoring teacher evaluation data to make professional development decisions changes the evaluation from a summative process to a formative one. In addition, the principal might also work with the instructional support team to ensure that the support being provided to teachers is high quality. Specifically, principals might focus on incorporating the use of student work into the feedback that is being provided to teachers. During the post conference process, evaluators might generate student evidence from observations of other lessons, from when they were in the classroom, or by trying the suggestions they plan to make in the classrooms themselves before providing the feedback in the post conference. Generating evidence for how the suggestion impacts students might be particularly useful for increasing perceptions of fairness when the evaluator is evaluating grade levels where they have little or no teaching experience.

By shifting the focus from evaluation for accountability to evaluation for teacher support, the principal may also be able to alter the perceptions of his/her beliefs about the evaluation process. To make the shift to a support centric process, the principal will have to generate structures and put focus on the evaluation process that emphasizes its usefulness as an instructional tool. By implementing a support centric evaluation process, principals may be able to alter how teachers perceive the evaluation process within their schools. They may be able to limit the amount of stress that teachers feel as a result of the evaluation, and help teachers

maintain or strengthen their teacher self-efficacy. A support centric evaluation model may also aid in assisting teachers in making instructional changes that enhance student learning.

Implications for District and State Leaders

District leaders should consider identifying schools that have generated support centric evaluation models. By identifying schools that are support centric in their evaluation practices, district leaders may be able to identify those structures and roles that are most useful to generating a school culture that supports teacher self-efficacy and results in positive changes to instructional practice. In addition, districts may consider the supports principals might need in order to shift to a support centric evaluation model. Districts might also consider the autonomy principals might need in order to generate a support centric evaluation process. District initiated structures and formal leadership roles may need to be reimaged to assist in creating a support centric evaluation system. District leaders might also consider ways to support the leadership skills needed to maintain a support centered evaluation process.

Tennessee state leaders should also consider identifying those schools and districts that have high functioning support centric evaluation systems. By leveraging the knowledge and skills that these school and district leaders possess, state leaders may be able to generate replicable models that would support and enhance the state adopted evaluation system. By having these school and district leaders share their experiences, state leaders might also be able to encourage more principals and districts to move to a support centric model for evaluation.

Recommendations for Evaluation Research

Further research might explore how schools move along a continuum from a procedural implementation, where the evaluation process is just a task to be completed, to making it an embedded part of school culture that generates positivity and growth. Additional research might

also explore the correlation between implementation practices and teacher self-efficacy. To do this, researchers might utilize the characteristics of implementation survey as the starting point for a valid and reliable instrument that would measure a school's evaluation culture.

In order to fully understand how the characteristics of evaluation implementation influence teacher self-efficacy recommendations for future research include a qualitative study utilizing teachers from across the state to see whether the findings in this study hold true for a broader population. Replicating this study with more participants could add more clarification to the ideas presented here. One of the challenges in this study was ensuring the safety of participants against retribution for sharing their evaluation experiences. Considering strategies for maximizing participants' ability to conceal their identities may also be beneficial in the design phase of such a study.

Recommendations for Efficacy Research

There has been some debate in the literature around how malleable efficacy is over time. The findings from the quantitative research seem to suggest that while teachers' overall self-efficacy may not vary much over time, the make-up of those beliefs does seem to shift and change. Additional research might explore how often and to what extent changes to teacher self-efficacy occur over time. Future studies might also explore the connection between these changes and actual changes in teacher practice. Exploring whether or not changes in efficacy over time are correlated to changes in teacher performance may provide more insight into the importance of teacher self-efficacy as a construct.

The exploration of affective states as a means of efficacy development has been largely absent in the literature. The findings from this study seem to suggest that the teacher's response to the stresses of evaluation influenced the power that was given to the other sources of efficacy

the evaluation process generated. Further research might explore the relationship between stress response and the other three sources of efficacy.

In addition to discussing the evaluation process, all of the teachers brought up student data as an influencing factor to their beliefs about their instruction.

Additional research might explore the role that data plays in influencing teachers' beliefs about their instruction.

Summary

Findings from this study indicated that the characteristics of evaluation implementation influenced teacher self-efficacy in a variety of ways. The TEAM evaluation system has the potential to generate positive or negative sources of efficacy. This means that in order for the evaluation process to be most beneficial to teachers and students, schools and districts need to pay close attention to how they are using the evaluation model within their schools. The actions of school and district leaders in how they communicate, engage in, and support the evaluation process have an impact on how the efficacy sources are generated and used. However, despite differences in the implementation and the impact on the teacher's efficacy, all of the teachers interviewed reported that they had seen positive changes to their instructional practice as a result of the evaluation. For a state that is trying to improve outcomes for students, this is good news.

In order to continue to strengthen instructional practice and improve outcomes for students, principals should create support centric evaluation models in order to better support teacher self-efficacy and limit the amount of stress teachers feel as a result of the evaluation process. By generating a support centric school culture with the TEAM system as the central component, school and district leaders can leverage the system to support and improve teacher efficacy, and improve teacher's perceptions of the evaluation process. In light of the correlation

between teacher efficacy and student achievement, having a way to support and increase teacher efficacy provides school leaders with a tool they can use to improve outcomes for students. For students in Tennessee, the continued use of the TEAM evaluation system, coupled with an intensive effort to ensure a support centric model for evaluation implementation practices could mean continued improvement in student outcomes over time.

References

- Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education*, 17, 86-95.
- Anderson, T. & Kanuka, H., (2003). *E-Research: Methods, strategies, and issues*. Boston, MA: Allyn & Bacon.
- Anfara, V., Brown, K., & Mangione, T., (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher*, 31(7). 28-38.
- Armor, D. Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., & Zellman, G., (1976). *Analysis of the school preferred reading program in selected Los Angeles minority schools*. (Report No. R-2007-LAUDS). Santa Monica, CA: The Rand Corporation. (ERIC Document Reproduction Service No. ED 130 243).
- Ashton, P. T., & Webb, R. B., (1986). *Making a difference: Teachers' sense of efficacy and achievement*. New York: Longman.
- Ashton, P.T., Webb, R. B., & Doda, N., (1983). *A study of teacher's sense of efficacy. Final Report, Executive Summary*. National Institute of Education.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*. 37(2), 122-147.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*. 84(2), 191-215.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*. 28(2), 117-148.
- Bandura, A. (1995). *Self-efficacy in changing societies*. New York, New York: Cambridge University Press.

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, New York: W. H. Freeman and Company.
- Bandura, A. & Adams, N., (1977). Analysis of self-efficacy theory of behavioral change. *Cognitive Therapy and Research*. 1(4), 287-310.
- Berman, P., McLaughlin, M., Bass, G. Pauley, E., & Zellman, G., (1977). *Federal programs supporting educational change. Vol: 7 Factors affecting implementation and continuation*. Santa Monica, CA: The Rand Corporation. (ERIC Document Reproduction Service No. ED 140 432).
- Bird, T. D. & Little, J. W., (1985). *Instructional Leadership in Eight Secondary Schools*. Boulder, CO: Center for Action Research, Inc.
- Boote, D. & Beile, P., (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, 34(6). 3-15.
- Brinson, D. & Steiner, L., (2007). *Building collective efficacy: How leaders inspire teachers to achieve*. Washington, D.C.: Center for Comprehensive School Reform and Improvement.
- Brookover, W.B., Schweitzer, J.H., Schneider, J.M., Beady, C.H., Flood, P., & Wisenbaker, J.M., (1978). Elementary school social climate and school achievement. *American Educational Research Journal*. 15, 301-318.
- Bryant, C. B. (2013). *Perceptions of Tennessee School Principals About the Tennessee Educator Acceleration Model*. Dissertation submitted to East Tennessee State University.
- Cantrell, S. C., Almasi, J. F., Carter, J. C. & Rintamaa, M., (2012). Reading intervention

- in middle and high schools: Implementation fidelity, teacher efficacy, and student achievement. *Reading Psychology*, 34(1), 26-58.
- Center on Educational Policy. (August 2007). Implementing the No Child Left Behind requirements: A report in the series from capital to the classroom: Year 5 of the No Child Left Behind Act. Retrieved from <http://www.cepd.org/document/docWindow.cfm?fuseaction=document.viewDocument&documentid=222&documentFormatId=4087>
- Celik, K. (2013). The relationship between individual innovativeness and self-efficacy levels of student teachers. *International Journal of Scientific Research in Education*, 6(1), 56-67.
- Ceylandag, F. (2009). *Teacher Self-Efficacy Beliefs Toward Measurement and Evaluation Practices*. Thesis submitted to the Graduate School of Social Sciences of Middle East Technical University.
- Chong, W. H. & Kong, C. A., (2012). Teacher collaborative learning and teacher self-efficacy: The case of lesson study. *The Journal of Experimental Education*, 80(3), 263-283.
- Colby S. A., Bradshaw, L. K., & Joyner, R. L., (2002). *Teacher Evaluation: A Review of Literature*. Paper presented for the Annual Meeting of the American Educational Research Association.
- Conly, S. & Glasman, N. S., (2008). Fear, the school organization, and teacher evaluation. *Educational Policy*, 22, 63-85.
- Conner, C. M. (2013). Commentary on two teacher evaluation systems: Moving towards a

- shared understanding of effective teaching. *School Psychology Quarterly*. 28(4), 342-346.
- Creswell, J. W. (1994). *Research Design: Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage.
- Danielson, C. (2012, November). Observing classroom practice. *Educational Leadership*. 32-37.
- DeChenne, S., Enochs, L., & Needham, M., (2012). Science, technology, engineering, and mathematics graduate teaching assistants teaching self-efficacy. *Journal of the Scholarship of Teaching and Learning*, 12(4), 102-123.
- Delvaux, E., Vanhoof, J., Tuytens, M., Vekemen, E., Devos, G., & Van Petegem, P., (2013). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Education*, 36, 1-11.
- Dellinger, A., Bobbett, J., Olivier, D., & Ellett, C., (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-S. *Teaching and Teacher Education*. 24, 751-766.
- Dembo, M. & Gibson, S., (1985). Teachers' sense of efficacy: An important factor in school improvement. *The Elementary School Journal*, 86(2), 173-184.
- DiClemente, C. C. (1981). Self-efficacy and smoking cessation maintenance: A preliminary report. *Cognitive Therapy and Research*, 5(2), 175-187.
- Educational Partners, Inc. (2007). *Research Brief: Teacher Evaluation*.
<http://www.educationpartnerships.org/>
- Gay, L.R., Mills, G.E., & Airasian, P., (2009). Educational research: Competencies for analysis and applications. Upper Saddle River, NJ: Pearson Education, Inc.italics

- Gibson, Sherri & Dembo, Myron H., (1984). Teacher efficacy: Construct validation. *Journal of Educational Psychology*. 76(4), 569-582.
- Glaser, B. & Strauss, A., (1967). *The Discovery of Grounded Theory*. Hawthorne, NY: Aldine de Gruyter, Inc.
- Goddard, R., Hoy, W., & Woolfolk Hoy, A., (2004). Collective efficacy beliefs: Theoretical developments, empirical evidence, and future directions. *Educational Researcher*. 33(3), 3-13.
- Goddard, Y., Goddard, R., & Tschannen-Moran, M., (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record*. 109(4), 877-896.
- Gravetter, F. & Wallnau, L., (2011). *Essentials of statistics for the behavioral sciences*. Belmont, CA: Wadsworth
- Guo, Y., Conner, C. M., Yang, Y., Roehrig, A. D., & Morrison, F. J., (2012). The effects of teacher qualification, teacher self-efficacy, and classroom practices on fifth grade literacy outcomes. *The Elementary School Journal*, 113(2), 3-24.
- Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, 4(1), 63-69.
- Guskey, T. R., & Passaro, P. D., (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, 31(3), 627-643.
- Harris, Alma (2008). Leading innovation and change: Knowledge creation by schools for schools. *European Journal of Education*. 43(2), 219-228.
- Herriott, R. E., & Firestone, W. A., (1983). Multisite qualitative policy research:

- Optimizing description and generalizability. *Educational Researcher*. 12(2), 14-19.
- Hesse-Biber, S., & Leavy, P., (2008). *Handbook of Emergent Methods*. New York, NY: The Guilford Press.
- Holt, B. (2011). *An Exploratory Study of Project Lead the Way Secondary Engineering Educators' Self-Efficacy*. Dissertation submitted to the Faculty of Virginia Polytechnic Institute and State University.
- Johnson, B. E. (2011). The speed and accuracy of voice recognition software-assisted transcription versus the listen and type method: A research note. *Qualitative Research*, 11(1), 91-97.
- Karimvand, P. (2011). The nexus between Iranian teachers' self-efficacy, teaching experience, and gender. *English Language Teaching*, 4(3), 171-183.
- Kimball, S. M., & Milanowski, A., (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, 45(1), 34-70.
- Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordan, K. A., (2010). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise?. *Educational Psychology Review*, 23, 21-43.
- Knox County Schools. (2013, December). *Knox County Teacher Survey*.
http://agenda.knoxschools.org/docs/2014/MID/20140121_269/2195_Teacher%20Survey%20Comments%20December%202013.pdfknoxschools.org
- Labone, E. (2004). Teacher efficacy: Maturing the construct through research in alternative paradigms. *Teaching and Teacher Education*, 20, 341-359.
- Lent, R., & Hackett, G., (1987). Career self-efficacy: Empirical status and future

- directions. *Journal of Vocational Behavior*. 30, 347-382.
- McLaughlin, M. W., & Pfeifer, S. R., (1986). *Teacher evaluation: Learning for improvement and accountability and case studies*. Stanford Education Policy Institute: Stanford University, CA.
- Michigan Association for School Administrators, (2013). *Tennessee Struggles to Get Teacher Evaluation Right*. <http://gomasa.org/news/tennessee-struggles-get-teacher-evaluation-right>
- Milanowski, A. (2011). Validity research on teacher evaluation systems based on the framework for teaching. Paper presented for the American Education Research Association conference.
- Merriam, S. B. (2009). *Qualitative Research: A Guide to Design and Implementation*. San Francisco, CA: Jossey-Bass.
- Moore, W.P., & Esselman, M. E., (1992). *Teacher efficacy, empowerment, and a focused instructional climate: Does student achievement benefit?* A paper presented for the American Educational Research Association.
- Moran, R. M. (2013). *Tennessee Educator Acceleration Model: Teacher Perceptions of One Policy Implementation*. A dissertation presented for the doctor of philosophy degree: University of Tennessee.
- Morse, J. (2003) Principals of mixed methods. In A. Tashakkori & C. Teddlie (Eds.) *Handbook of mixed methods in social and behavioral research*. (pp. 189-206). Thousand Oaks, CA: Sage.
- National Institute for Excellence in Teaching. (2012). *TAP Research Summary*. http://www.tapsystem.org/publications/tap_research_summary_0210.pdf

New Teacher Project (2009). *The Widget Effect*.

<http://widgeteffect.org/downloads/TheWidgetEffect.pdf>

Olivier, D. F. (2001). *Teacher Personal and School Culture Characteristics in Effective Schools: Toward a Model of a Professional Learning Community*. Unpublished doctoral dissertation. Louisiana State University.

Patton, M. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage Publications, Inc.

Patton, M. (2002). *Qualitative evaluation and research methods* (3rd ed.). Newbury Park, CA: Sage.

Ross, J. (1992). Teacher efficacy and the effect of coaching on student achievement. *Canadian Journal of Education* 17(1), 51-65.

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

Schacter, J., & Thum, Y. M., (2004). Paying for high- and low-quality teaching. *Economics of Education Review*, 23, 411-430.

Simoncini, K., Lasen, M., & Rocco, S. (2014). Professional dialogue, reflective practice and teacher research: Engaging early childhood pre-service teachers in collegial dialogue about curriculum innovation. *Australian Journal of Teacher Education*, 39(1), 27-44.

Stajkovic, A., & Luthans, F., (1998). Social cognitive theory and self-efficacy: Going beyond traditional motivational and behavioral approaches. *Organizational Dynamics*. Spring, 62-74.

Stajkovic, A., & Luthans, F., (1998). Self-efficacy and work related performance: A meta-analysis. *Psychological Bulletin*. 124(2), 240-261.

- Strauss, A., & Corbin, J., (1990). *Basics of Qualitative Research*. Newbury Park, CA: Sage Publications, Inc.
- Tashakorri , A., & Teddlie, C., (1998). *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage.
- Tennessee Department of Education. (2013). Tennessee Educator Acceleration Model. <http://team-tn.org>
- Tennessee Government. (2013, November 7). *Tennessee Students the Fastest Improving in the Nation*. <http://news.tn.gov/node/11644>
- The Danielson Group. (2013). *The Framework for Teaching Evaluation Instrument*. <http://www.danielsongroup.org/userfiles/files/downloads/2013EvaluationInstrument.pdf>
- Tschannen-Moran, M., Woolfolk Hoy, A. & Hoy, W. K., (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, 68(2), 202-248.
- Tschannen-Moran, M., & Johnson, D., (2011). Exploring literacy teacher' self-efficacy beliefs: potential sources at play. *Teaching and Teacher Education*, 27, 751-761.
- Tschannen-Moran, M., & McMaster, P., (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *The Elementary School Journal*, 110(2), 228-245.
- Tschannen- Moran, M., & Woolfolk Hoy, A., (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Tschannen-Moran, M., & Woolfolk Hoy, A., (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education*, 23, 944-956.

- Weber, J. R. (1987). *Teacher Evaluation as a Strategy for Improving Instruction: Synthesis of Literature*. Elmherst, IL: North Central Regional Educational Lab.
- Weinberg, R., Gould, D., & Jackson, A., (1979). Expectations and performance: An empirical test of Bandura's self-efficacy theory. *Journal of Sport Psychology*, 1, 320-331.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D., (2009). *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness*. The New Teacher Project.
- Wheatley, K. F. (2002). The potential benefits of teacher efficacy doubts for educational reform. *Teaching and Teacher Education*, 18, 5-22.
- Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education*, 21, 747-766.
- Woolfolk, A., & Hoy, W., (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology*, 82, 81-91.
- Woolfolk Hoy, A., & 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.
- Yin, Robert K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.

Appendix

Teacher Efficacy Belief Scale

Self-Form

Directions: This survey requests that you make judgments about the **strength of your personal beliefs** in your capabilities to organize and successfully carry out teaching tasks in your school. In assessing the strengths of your personal beliefs about each tasks, consider your abilities within the context of your current school. Consider job roles and responsibilities, available resources and support, current policies, help from colleagues and so on. For each item, use the scale provided below and circle one of the corresponding numbers that best reflects the strength of your personal beliefs about your capabilities to accomplish each teaching task.

STRENGTH OF BELIEFS SCALE: 1 = *Weak Beliefs (WB)* in my capabilities:

2 = *Somewhat Strong Beliefs (SSB)* in my capabilities:

3 = *Strong Beliefs (SB)* in my capabilities:

4 = *Very Strong Beliefs (VSB)* in my capabilities:

In my present teaching situation, the strength of my <i>personal beliefs in my capabilities to . . .</i>		WB	SSB	SB	
		VSB			
		1	2	3	4
1.	plan activities that accommodate the range of individual differences among students...	1	2	3	4
2.	plan evaluation procedures that accommodate individual differences among students...	1	2	3	4
3.	use allocated time for activities that maximize learning...	1	2	3	4
4.	effectively manage routines and procedures for learning tasks...	1	2	3	4
5.	clarify directions for learning routines...	1	2	3	4
6.	maintain high levels of student engagement in learning tasks...	1	2	3	4
7.	redirect students who are persistently off task...	1	2	3	4
8.	maintain a classroom climate of courtesy and respect...	1	2	3	4
9.	maintain a classroom climate that is fair and impartial...	1	2	3	4
10.	communicate to students the specific learning outcomes of the lesson...	1	2	3	4
11.	communicate to students the purpose and/or importance of learning tasks...	1	2	3	4
12.	implement teaching methods at an appropriate pace to	1	2	3	4

accommodate differences among students...

13.	utilize teaching aids and learning materials that accommodate individual differences among students...	1	2	3	4
14.	provide students with opportunities to learn at more than one cognitive and/or performance level...	1	2	3	4
15.	communicate to students content knowledge that is accurate and logical...	1	2	3	4
16.	clarify student misunderstandings or difficulties in learning...	1	2	3	4
17.	provide students with specific feedback about their learning...	1	2	3	4
18.	provide students with suggestions for improving learning...	1	2	3	4
19.	actively involve students in developing concepts...	1	2	3	4
20.	solicit a variety of questions throughout the lesson that enable higher order thinking...	1	2	3	4
21.	actively involve students in critical analysis and/or problem solving...	1	2	3	4
22.	monitor students involvement during learning tasks...	1	2	3	4
23.	adjust teaching and learning activities as needed...	1	2	3	4
24.	manage student discipline/behavior...	1	2	3	4
25.	involve students in developing higher order thinking skills...	1	2	3	4
26.	motivate students to perform to their fullest potential...	1	2	3	4
27.	provide a learning environment that accommodates students with special needs...	1	2	3	4
28.	improve the academic performance of students, including those with learning abilities...	1	2	3	4
29.	provide a positive influence on the academic development of students...	1	2	3	4
30.	maintain a classroom environment in which students work cooperatively...	1	2	3	4

Teacher Survey Items for Characteristics of Evaluation Implementation

The following statements are about your perceptions of a variety of school factors. Please indicate the extent to which you agree with each of the following statements along the following scale:

1- Strongly disagree

2- Disagree

3- Agree

4- Strongly agree

N/A- I choose not to provide an answer for this statement.

- | | | | | | |
|---|---|---|---|---|-----|
| 1. My evaluators are qualified to accurately assess my instruction. | 1 | 2 | 3 | 4 | N/A |
| 2. My principal views the evaluation process as a way to help teachers improve instruction. | 1 | 2 | 3 | 4 | N/A |
| 3. Teachers and evaluators work together to improve instruction in my school. | 1 | 2 | 3 | 4 | N/A |
| 4. The feedback provided to me in the evaluation process is meaningful to me. | 1 | 2 | 3 | 4 | N/A |
| 5. The TEAM/TAP rubric is an effective evaluation tool. | 1 | 2 | 3 | 4 | N/A |
| 6. My principals utilizes the rubric to provide me with frequent feedback about my performance. | 1 | 2 | 3 | 4 | N/A |
| 7. My evaluation scores are an accurate reflection of my teaching ability. | 1 | 2 | 3 | 4 | N/A |
| 8. The relationship between teachers and evaluators is positive. | 1 | 2 | 3 | 4 | N/A |
| 9. I utilize the suggestions made by my evaluators in order to grow professionally. | 1 | 2 | 3 | 4 | N/A |
| 10. My evaluator provides me with specific suggestions for improvement. | 1 | 2 | 3 | 4 | N/A |

Teacher Interview Protocol (Original)

1. Tell me about the evaluation process. What happens in a typical evaluation?
2. In the post conference I observed, your principal said, “feedback example.” What did that make you think? How did it make you feel?
3. In the post conference I observed, you were refined in “indicator/descriptor.” How will that impact your future instruction? What about the evaluation process influenced that decision?
4. In the post conference I observed, you were reinforced in “indicator/descriptor.” How will that impact your future instruction? What about the evaluation process influenced that decision?
5. What did you think about your ability to “teaching task from quantitative data that changed” before this evaluation? How has that changed? What led to that change?
6. I noticed that you also changed your belief about your ability to “teaching task from quantitative data that changed.” What led to that change?
7. How beneficial was it for your evaluator to “follow-up activity” following your post conference?
8. How would you feel if that follow-up had not been provided?

Teacher Interview Protocol (Revised)

1. Tell me about the evaluation process. What happens in a typical evaluation?
2. What feedback did you receive through that process? How did you use that feedback?
3. What influenced your decision to use the feedback/not use the feedback that way?
4. What do you find most beneficial about the evaluation process? What do you find the least beneficial about the evaluation process?
5. What did you think about your ability to “teaching task from quantitative data that changed” before this evaluation? How has that changed? What led to that change?
6. I noticed that you also changed your belief about your ability to “teaching task from quantitative data that changed.” What led to that change?
7. Do you believe that the teacher’s relationship with the evaluator influences what the teacher gets out of the process?

Principal Interview Protocol

School:

Principal:

Number of Years as Principal:

1. Describe the role teacher evaluation plays in your school.
2. What do you perceive to be the purpose of teacher evaluation?
3. How do you use the information you gain from the evaluations?
4. What do your teachers think of the evaluation system?
5. How do you ensure that your evaluations are fair and valid?
6. How would teachers in your building describe the fairness and validity of the evaluation process? Why?
7. How do you prepare for a post conference?
8. What do you think makes a post conference successful?
9. If you had to capture the relationship between the evaluator and the teacher in one word, what would it be and why?

Permission Letter for TEBS-Self



Department of Educational Foundations
and Leadership
P.O. Box 43091
Lafayette, LA 70504-3091

June 2, 2014

Liz Norton
Graduate Student
Educational Leadership and Policy Studies
University of Tennessee
317 East Glenwood Ave.
Knoxville, TN 37917

Dear Ms. Norton:

This correspondence is to grant permission to utilize the *Teacher Efficacy Beliefs Scale-Self Form* (TEBS-S) as one of your instruments for data collection for your doctoral study through the University of Tennessee. I believe your research exploring *how Tennessee's evaluation model influences teacher self-efficacy* will contribute to both educational evaluation and self-efficacy literature.

This permission letter allows use of the TEBS-S through either paper/pencil or online administration.

Upon completion of your study, I would be interested in learning about your completed study and would welcome the opportunity to receive an electronic version of your completed dissertation research.

Thank you for your interest in my research and the TEBS-S. Should you require any additional information, please feel free to contact me.

Sincerely,

Dianne F. Olivier

Dianne F. Olivier, Ph. D.
Associate Professor
Joan D. and Alexander S. Haig/BORSF Professor
Department of Educational Foundations and Leadership
College of Education
University of Louisiana at Lafayette
P.O. Box 43091
Lafayette, LA 70504-3091
(337) 482-6408 (Office) dolivier@louisiana.edu

Implementation Characteristics Survey Results

Frequency Count School A

	1	2	3	4	N/A
1. My evaluators are qualified to accurately assess my instruction.	0	0	6	1	0
2. My principal views the evaluation process as a way to help teachers improve instruction.	0	1	2	4	0
3. Teachers and evaluators work together to improve instruction in my school.	0	2	2	2	0
4. The feedback provided to me in the evaluation process is meaningful to me.	0	1	5	1	0
5. The TEAM/TAP rubric is an effective evaluation tool.	3	1	3	0	0
6. My principals utilizes the rubric to provide me with frequent feedback about my performance.	0	0	6	1	0
7. My evaluation scores are an accurate reflection of my teaching ability.	1	1	3	1	1
8. The relationship between teachers and evaluators is positive.	1	0	5	1	0
9. I utilize the suggestions made by my evaluators in order to grow professionally.	1	0	3	2	1
10. My evaluator provides me with specific suggestions for improvement	0	1	4	2	0

Frequency Count School B

	1	2	3	4	N/A
1. My evaluators are qualified to accurately assess my instruction.	0	1	7	1	0
2. My principal views the evaluation process as a way to help teachers improve instruction.	0	1	5	1	2
3. Teachers and evaluators work together to improve instruction in my school.	0	3	4	1	1
4. The feedback provided to me in the evaluation process is meaningful to me.	0	2	4	3	0
5. The TEAM/TAP rubric is an effective evaluation tool.	4	3	2	0	0
6. My principals utilizes the rubric to provide me with frequent feedback about my performance.	0	3	3	2	1
7. My evaluation scores are an accurate reflection of my teaching ability.	2	1	6	0	0
8. The relationship between teachers and evaluators is positive.	1	1	5	2	0
9. I utilize the suggestions made by my evaluators in order to grow professionally.	0	0	6	3	0
10. My evaluator provides me with specific suggestions for improvement	0	2	2	5	0

Frequency Count School C

	1	2	3	4	N/A
1. My evaluators are qualified to accurately assess my instruction.	0	0	2	2	0
2. My principal views the evaluation process as a way to help teachers improve instruction.	0	0	1	3	0
3. Teachers and evaluators work together to improve instruction in my school.	0	1	0	3	0
4. The feedback provided to me in the evaluation process is meaningful to me.	0	0	2	2	0
5. The TEAM/TAP rubric is an effective evaluation tool.	0	0	1	2	1
6. My principals utilizes the rubric to provide me with frequent feedback about my performance.	0	1	0	3	0
7. My evaluation scores are an accurate reflection of my teaching ability.	0	0	1	2	1
8. The relationship between teachers and evaluators is positive.	0	0	0	4	0
9. I utilize the suggestions made by my evaluators in order to grow professionally.	0	0	0	4	0
10. My evaluator provides me with specific suggestions for improvement	0	0	2	2	0

Vita

Elizabeth Norton was born in Hammond, Louisiana. In 1997 she moved to Kingston, Tennessee and graduated from Roane County High School in 1999. She obtained a Bachelor's degree in Elementary Education in 2003 and a Master's degree in Curriculum and Instruction in 2008, both from Tennessee Technological University. She has worked with the Knox County Schools since 2003 in various capacities including teacher, instructional coach, and master teacher. In 2015, she completed a Doctor of Philosophy Degree in Educational Leadership and Policy Studies at the University of Tennessee. She currently works as an Executive Master Teacher supporting 12 Knox County schools in their implementation of TAP, The System for Teacher and Student Advancement.