THE MEANING AND MEASURE OF TEACHER

SELF-EFFICACY FOR EFFECTIVE CLASSROOM TEACHING PRACTICES

by

Colleen Monica Kuusinen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Education and Psychology)
in The University of Michigan
2016

Doctoral Committee:

Research Professor Stuart A. Karabenick, Chair
Professor Tabbye M. Chavous
Clinical Assistant Professor Debi Khasnabis
Assistant Professor Fani Lauermann, University of Bonn
Professor Allison M. Ryan

DEDICATION

This dissertation is dedicated to my mother, Marcela Guardia Vargas for teaching me to love psychology and to never give up.

ACKNOWLEDGEMENTS

There are so many people who, whether realizing it or not, helped me to produce this dissertation and finish graduate school with a smile on my face. First, my advisor, Stuart Karabenick. It took me a few years, Stuart, to think like a researcher, but I got there, and our conversations were a big part of that. Thank you for giving me the freedom to pursue my interests, design my own studies, and never doubting that I would get where I needed to be in the end. My committee, Allison Ryan, Fani Lauermann, Debi Khasnabis, and Tabbye Chavous. What an honor to work with outstanding scholars with such diverse areas of interest. You are role models for women in academia in so many different ways. To David Cohen and Matthew Ronfeldt, who read early versions of my work and gave me the confidence that I could write like an academic and that what I had to say was worth writing about; to Deborah Ball, who taught me crucial, paradigm-changing ideas about teaching through simply modeling expert teaching; to Eric Fretz, who made being a TA a pleasure; to Janie Knieper and the CPEP community that lifts everyone up and models equitable education: what a privilege. I have been blessed.

My friends in Ann Arbor, both in and out of graduate school: Seema Jolly, Noah Neidlinger, Kathryn Curtis, Ryan McKenzie, Darian McKenzie, Lisa Arthur, Kate Nestor, Matthew Hamburg, Stephanie Moore, Jevon Moore, Becky Christensen, Lotus Yu, Peter Sherer, Laura Urteaga-Fuentes, Kara Makara Fuller, Keith Fuller. The laughter, the dancing, the betting, the potlucks, the grilled pizza parties, the costuming, the camping, the

games, the tears, the road trips. You were my family here. Your friendship has changed me forever.

To my mother, who within three weeks of me defending this dissertation finally achieved her life dream to become a licensed clinical psychologist. At the young age of 72. Our long ago, late-night conversations are what inspired my curiosity about people and why they act the way they do; our recent, weekly conversations have reminded me that our families, however we define them, make us stronger than we ever thought we could be. You are going to be happy doing what you love. May we all be so lucky.

Table of Contents

Dedication ii

Acknowledgements iii

List of Tables vi

List of Figures vii

Abstract viii

CHAPTER 1: INTRODUCTION 1

CHAPTER 2: REVIEW OF THE RELEVANT LITERATURE 5

Teacher efficacy as implicit theory of student motivation and performance 9
Teacher self-efficacy reconceptualized: The present era 18
Domains of effective instruction 34
Measuring multiple domains of teacher self-efficacy 46
The role of teachers' experience and context 49
Conclusion and looking ahead 52

CHAPTER 3: A META-ANALYTIC REVIEW OF TEACHER SELF-EFFICACY AND EFFECTIVE CLASSROOM TEACHING PRACTICES 72

Multi-dimensionality of teacher self-efficacy and teaching behavior 73
Three domains of effective classroom teaching 75
Prior reviews of teacher self-efficacy and teaching quality 79
Importance of item construction 80
Method 82
Results 86
Discussion 92

CHAPTER 4: THE MEANING OF TEACHERS' SELF-EFFICACY BELIEFS: PERSPECTIVES OF MIDDLE AND HIGH SCHOOL TEACHERS 113

Teacher self-efficacy research and the need for qualitative analyses 115
The role of students 117
The role of teachers' experience 118
Method 121
Results 125
Discussion 157

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS 184

List of Tables

Table 1. Classroom Management and Discipline Scale (Emmer & Hickman, 1991) 25

Table 2. Structured Classroom Management indicators synthesized from literature 37

Table 3. Supportive Classroom Climate indicators Synthesized from Literature 39

Table 4. Cognitive Demand indicators Synthesized from Literature 44

Table 5. Indicators of effective classroom teaching by domain 97

Table 6. TSE measures used in studies 98

Table 7. Observational measures used in studies 98

Table 8. Domain matching in coded studies 99

Table 9. All studies meeting eligibility criteria 100

Table 10. Summary of effect sizes for full model and by domains of teaching practice .. 102

Table 11. Summary of moderators for Structured Management model (k=21) 102

Table 12. Summary of moderators for Supportive Climate model (k=14) 102

Table 13. Summary of moderators for Cognitive Activation model (k=18) 103

Table 14. Demographic information 165

Table 15. Selected Codes 166

Table 16: Multiple meanings of TSE for student outcomes 167

Table 17: Multiple meanings of TSE for strategies 169

List of Figures

Figure 1. Bandura (1997) model for efficacy beliefs and outcome expectations with performance attainment added 7

Figure 2. Ashton's model of teacher efficacy (1983) 11

Figure 3. Portion of process model showing formation of teacher self-efficacy beliefs (Tschannen-Moran et al., 1998) 20

Figure 4. Conceptual Model for Teaching as Interaction (Cohen, Raudenbush & Ball, 2003) 31

Figure 5. Moderator coding by teacher self-efficacy scale 103

Figure 6. Domain coding by version and subscale of TSES 104

Figure 7. Synthesis forest plot for included studies 104

Figure 8. Application of SO code across subscales 170

Figure 9. Application of SO code to four Instructional Strategies Efficacy (ISE) items 170

Abstract

A substantial body of research has established the importance of teacher self-efficacy (TSE) for teachers' psychological well-being and general attitudes toward students. However, few studies conclusively link TSE with research-based practices for effective classroom teaching. Underlying this concern are two unresolved issues in TSE research: (a) conceptual ambiguity about the role of students in the teaching tasks upon which TSE beliefs are based, and (b) debate about the appropriate manner in which to measure TSE beliefs. This dissertation employs quantitative and qualitative methods to provide insight into these two important issues. The first study is a meta-analytic review of the relationship between TSE and three domains of effective classroom teaching synthesized from the literature: structured classroom management, supportive classroom climate, and cognitive activation. Included were potential moderators of this association related to the measurement of TSE (e.g., social cognitive theoretical orientation of measure, percent of items including a problem). Thirty-nine studies met inclusion criteria, which provided 133 effect sizes (Pearson r correlations) between TSE and observations of classroom teaching in any of the three domains. The overall effect size was small ($r = .17$), with a trend towards social cognitive measures having stronger effects. No moderators reached statistical significance at the .05 level. However, the inclusion of problems in the TSE measure was marginally greater ($p = .06$.)

The qualitative study used think aloud methodology to capture teachers' spontaneous self-efficacy judgments in response to items of a commonly used social cognitive measure of TSE, the Teachers' Sense of Efficacy Scale (Tschannen-Moran & Hoy, 2001). Responses from 23 secondary teachers in core academic subjects yielded multiple interpretations of survey items and, thus, multiple meanings of TSE beliefs. In addition, differences in meaning were observed among student, beginning, and experienced teachers. However, responses to problem-based TSE items were more consistent, representing beliefs in teacher capability to solve problems in the classroom through integration of knowledge of students with pedagogical knowledge. The discussion suggested ways in which a problem-based approached to measuring TSE could translate into practical implications for predicting teachers' use of effective classroom teaching practices and their professional development.

CHAPTER 1

INTRODUCTION

The identification and retention of quality teachers is an area of international concern. The teaching profession in the U.S., for example, is characterized by high turnover within the first five years of teaching (Cochran-Smith et al., 2012; Ingersoll, 2001). Researchers and practitioners have acknowledged the importance of teacher motivation, particularly their sense of self-efficacy, as a strong predictor of their job satisfaction and intention to stay in the profession (Klassen & Chiu, 2010; Skaalvik & Skaalvik, 2010). Representative of teachers' beliefs that they can accomplish teaching tasks at a desired level of competence in a specific context, teacher self-efficacy has also been associated with such teacher practices as greater persistence with students who are struggling (Allinder, 1995).

Unresolved, however, is whether teachers' sense of efficacy is related to use of specific teaching practices known to be effective for student learning. Conclusions about this relationship are limited because most empirical work in this area preceded the development of conceptually sound teacher self-efficacy measures and reliable classroom observation tools to assess research-based components of effective teaching (e.g., Evertson, Anderson, Anderson, & Brophy, 1980; Kounin, 1970). The development of reliable measures in the last 15 years such as the Teachers Sense of Efficacy scale (TSES; Tschannen-Moran & Hoy, 2001) and the Classroom Assessment Scoring System (CLASS; Pianta, Belsky, Vandergrift, Houts, & Morrison, 2008) allow more nuanced and reliable investigations into the relationship between teacher self-efficacy and effective teaching practice. Since teacher self-efficacy measures were not developed in tandem with

observational measures of effective teaching, however, they may reflect different foci of teaching and limit the predictive validity of such scales for effective teaching practice.

Moreover, a critical conceptual issue not sufficiently addressed in the literature to date is the meaning of teacher self-efficacy judgments with respect to students and student outcomes. Studies that have attempted to establish a relationship between teacher self-efficacy as assessed by the TSES and student achievement have yielded inconclusive results or weak relationships (Bruce, Esmonde, Ross, Dookie, & Beatty, 2010). Underlying this concern is the conceptual issue of whether teachers' sense of efficacy represents confidence in their capability to (a) have a direct influence on student learning, engagement and behavioral outcomes or (b) accomplish teaching tasks with the purpose of creating structured, supportive and cognitively demanding classroom environments despite challenges that may exist in the environment, or (c) both of these goals. In other words, it is unclear whether teachers' sense of efficacy is based on their perceived capability to achieve student outcomes or on their perceived capability to perform tasks that are likely to bring about student outcomes, i.e., effective teaching practices.

Problem Statement

Critical to addressing this issue is clarifying the meaning and implications of teacher's self-efficacy judgments, particularly in relation to effective teaching practices and student outcomes. As many different ways to assess teacher self-efficacy exist (e.g., efficacy for a student outcome, a teaching strategy, or to solve a problem of practice), it is unknown whether the approach to assessment influences the criteria teachers use to judge their self-efficacy, and thus the meaning of their self-efficacy judgments. Additionally, researchers

have suggested that the meaning of teachers' self-efficacy judgments, not just the strength of those beliefs, depend on such personal characteristics as their years of work or prior learning experiences (Tschannen-Moran, Hoy, & Hoy, 1998; Wheatley, 2005). Finally, as research has begun to converge on a multi-dimensional view of effective teaching practices defined by at least three dimensions (Brophy, 2006; Holzberger, Philipp, & Kunter, 2013; Klieme, 2012), it is critical to examine the extent that teachers' self-efficacy, itself multi-dimensional, is related to corresponding multi-dimensional teaching practices. Such a multi-pronged investigation would address recent calls to move teacher self-efficacy from theory to practice (Klassen, Durksen, & Tze, 2014).

Overview

This multiple manuscript dissertation explores the meaning, measurement and implications of teacher self-efficacy for effective classroom teaching. Specifically, Chapter 2 reviews the extant literature on teacher self-efficacy, describing important conceptual shifts in its meaning and measurement. The following two studies systematically investigate the empirical relationships between teacher efficacy and effective classroom teaching practices and explore the meaning of teacher self-efficacy TSE beliefs in relation to these practices, student outcomes, and teachers' context. Chapter 3 provides a meta-analysis that synthesizes evidence of the associations between teacher self-efficacy and observations of effective teaching practices, accounting for potential mediating effects of years of experience and measurement of TSE. Chapter 4 is a qualitative study that (a) summarizes the criteria that teachers use to judge their self-efficacy in response to the TSES, particularly in relation to student outcomes (b) explores the influence of item construction

on teacher thinking and (c) determines the overall content validity of the TSES items, particularly in relation to research-based strategies for effective instruction, from the perspective of teachers. Chapter 5 presents the implications of these studies for the conceptualization and assessment of TSE.

CHAPTER 2

REVIEW OF THE RELEVANT LITERATURE

Theoretical foundations of teacher self-efficacy

In the following section, I describe the social cognitive theory of self-efficacy and summarize different theoretical approaches to the study of teacher self-efficacy within this tradition. First I begin with describing the social cognitive theory of self-efficacy. Subsequently, three strands of teacher self-efficacy research are described, one based on teacher efficacy as an implicit theory of student motivation and performance and the latter focused on studying teacher self-efficacy as a multi-dimensional, task-specific set of efficacy beliefs.

Social cognitive theory of self-efficacy

There are four theorized sources of efficacy beliefs within social-cognitive theory (Bandura, 1977a, 1997; Tschannen-Moran et al., 1998). The first and most impactful are *mastery experiences*, those experiences which teachers believe to have achieved success in performing a teaching task. *Vicarious experiences* are those in which a teacher observes another successfully completing the teaching task; these experiences are most impactful on self-efficacy beliefs when the teacher both trusts the person performing the task and feels that they possess similar characteristics to that person. Similarly, *verbal persuasion* can also affect a teacher's self-efficacy beliefs if the person providing positive feedback to the

teacher is trusted or seen as more capable than the teacher. The final source of self-efficacy beliefs is *affective experiences*—feelings of anxiety, joy, pride, and despair can impact how individuals judge their capabilities.

Important for the present research, Bandura (1997) clearly distinguishes between efficacy beliefs, outcome expectancies, and markers of different performance levels. As depicted in Figure 1, efficacy beliefs are strong predictors of behavior while outcome expectancies are the physical, social and self-evaluative outcomes that result from performing a task at a desired level of competency. Central to self-efficacy theory, self-efficacy beliefs and outcome expectancies are beliefs about future functioning or outcomes, respectively. They are not, for example, retrospective beliefs about what led to success or failure on a specific task: those would be attributions, defined by attribution theory (Weiner, 1985). A performance attainment is an accomplishment specified by descriptive markers (e.g., academic work at an A-level of performance) while an outcome is the consequence of a performance (e.g., self-satisfaction and social approval).

Bandura (1997) points out that great conceptual confusion is created when performance markers are mistaken for outcome expectancies. For example, for the task of doing well on a mathematics test, getting an A on that test is not the outcome expectancy but the performance marker of having successfully completed that task. In other words, the specific task is actually completing a mathematics test with an A-level of performance. The A-level indicates the difficulty of the task against which individuals can judge their sense of efficacy. There are both conceptual and practical arguments to support this claim. Conceptually, behavior is distinct from the effect it produces, and effects (i.e., outcomes) can be used to influence behavior. If an A-level performance on a mathematic test were

construed as an outcome, then one is arguing that an A-level performance is needed to produce an A level performance (Bandura, 1997, p.23).

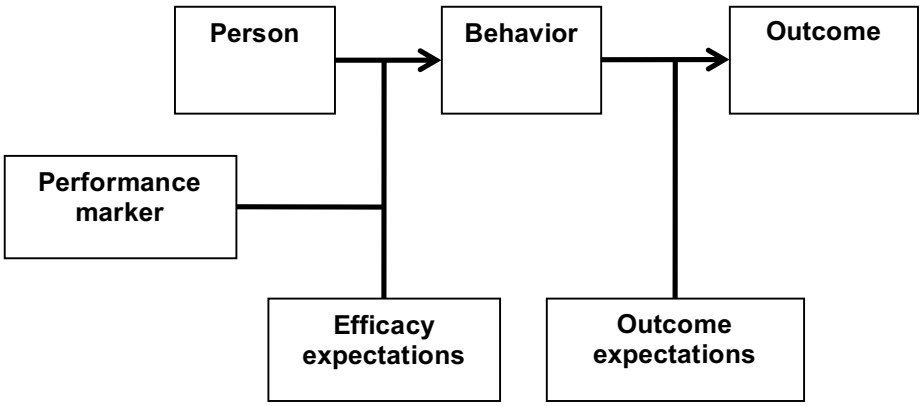

Figure 1. Bandura (1997) model for efficacy beliefs and outcome expectations with performance attainment added

Practically, without such an indicator of the target level of performance by which individuals can judge their sense of efficacy, implications of comparing ratings across individuals (who may have different desired levels of performance) is limited to general judgments of how confident individuals feel about particular behaviors, no matter the quality of the performance of those behaviors. Individuals with lower expectations for successful performance may feel more efficacious and would likely only persist until they reached their lower desired level of performance. This consideration is most salient when one wishes to use self-efficacy ratings to predict the quality of one's performance for a task, such as is the case when using self-efficacy ratings to predict the quality of one's teaching performance.

Finally, it is important to note that outcome expectancies are distinct from attributions for success or failure, and important distinction for understanding research conducted in this era. Though both constructs are related to success (i.e., mastery)

experiences, external attributions and outcome expectancies differ in temporality. Attributions are past-referent; they are how one explains success or failure *after* it has occurred (Weiner, 1985). Attributions can be made to internal factors, such as one's own effort and skill, or to external factors such as luck or the resources available to complete the task. Outcome expectancies are forward-looking; they are what one believes *will* happen *if* they are successful. Moreover, outcome expectancies do little to predict behavior, as one can expect certain outcomes from successful performance but feel incapable of performing the necessary functions to accomplish that task.

Supporting this conceptual and practical distinction between outcome expectancies and performance markers, the studies in this dissertation investigate on (a) what teachers base their sense of efficacy to teach and (b) in what manner and under what conditions is evidence of desired student outcomes construed as a performance marker for success in teaching tasks. More specifically, the following questions are explored:

- Is successful attainment of student outcomes—productive changes in student thinking, attitude or behavior—represented by the performance markers or can teachers perceive success even when they fail to achieve desired student outcomes?
- Does this criterion vary by which type of task they consider (e.g., management vs. instruction) or by how much work experience they have (e.g., novice versus experienced teacher)?
- How strongly is teachers' self-efficacy related to the quality of their teaching in different dimensions of research-based classroom teaching strategies?

The subsequent four sections describe three distinct ways in which teacher efficacy, outcome expectancies and student outcomes have been considered in the literature. The first section describes the original conceptualization of teacher efficacy as implicit beliefs about student motivation and performance. The second and third sections describe efforts by researchers to use psychometric studies of teacher efficacy scales to determine what constitutes outcome expectancies and efficacy beliefs. Fourth, the most current conceptualization of teacher self-efficacy grounded in self-efficacy theory is described. Through describing these four shifts in conceptualization of teacher self-efficacy, I show how a subtle change occurred in the conceptualization and measurement of teacher self-efficacy beliefs moving the field away from considering beliefs about student performance as an inherent part of teachers' self-efficacy beliefs to one based on evaluation of capability to implement strategies at an undefined level of performance.

Teacher efficacy as implicit theory of student motivation and performance

The identification and study of teacher efficacy originated with a series of empirical studies conducted by the RAND corporation (Armor et al., 1976; Berman, McLaughlin, Bass, Pauly, & Zellman, 1977). In a study of minority student achievement in reading, two items were found to strongly predict variation in Black students' achievement in reading ($r = .31$). The authors reported that these items assessed "the extent to which the teacher believes he or she has the capacity to produce an effect on the learning of students" (Armor et al., 1976, p. 23). In a footnote, the authors stated that the two items used were based on Rotter's (1966) model for generalized expectancies for internal versus external control of reinforcement, often called locus of control. Locus of control refers to the degree to which an individual expects that an outcome of their behavior is contingent either on their own

actions or personal characteristics versus chance, luck, fate or other external factors (Rotter, 1966). RAND item 1 assessed a teacher's belief about the potential impact of any teacher in the face of external influences of the home environment: *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.* Rand item 2 focused on teachers' beliefs in their personal capabilities: *If I really try hard, I can get through to even the most difficult or unmotivated students.* Taken together, these items indicate that efficacy is a belief that teachers can get through to students by influencing their motivation and performance.

As researchers began to interpret the meaning of teacher efficacy in more depth, the internal/external control orientation was overlooked in favor of a focus on the general vs. personal distinction implied by the use of "a teacher" in the first RAND item and "I" in the second RAND item. Among the first was Ashton (1983, 1984a; Ashton, Olejnik, Crocker, & McAuliffe, 1982; Ashton & Webb, 1986), who also developed a more comprehensive conceptual framework for teacher self-efficacy research. Ashton and colleagues offered two definitions for teacher' sense of efficacy. First, teachers' sense of efficacy was defined as an "implicit personal theory of student motivation" (1983, p. 2), according to which student performance is implied to result from students' greater motivation. In the publication resulting from this report, teacher efficacy was defined as "teachers' situation-specific expectation that they can help students learn. That expectation rest on assumptions of how much students are capable of learning what schools have to teach" (1984, p. 3). Thus in the latter definition, more widely cited, greater emphasis is placed on teacher efficacy as a belief about what teachers can do to help students learn rather than the

mechanisms by which students learn. Ashton grounded her model in Bandura's theory of self-efficacy, though later theorists tended to categorize her work (later with Webb) in the Rotter and RAND tradition or as "other" (Tschannen-Moran et al., 1998).

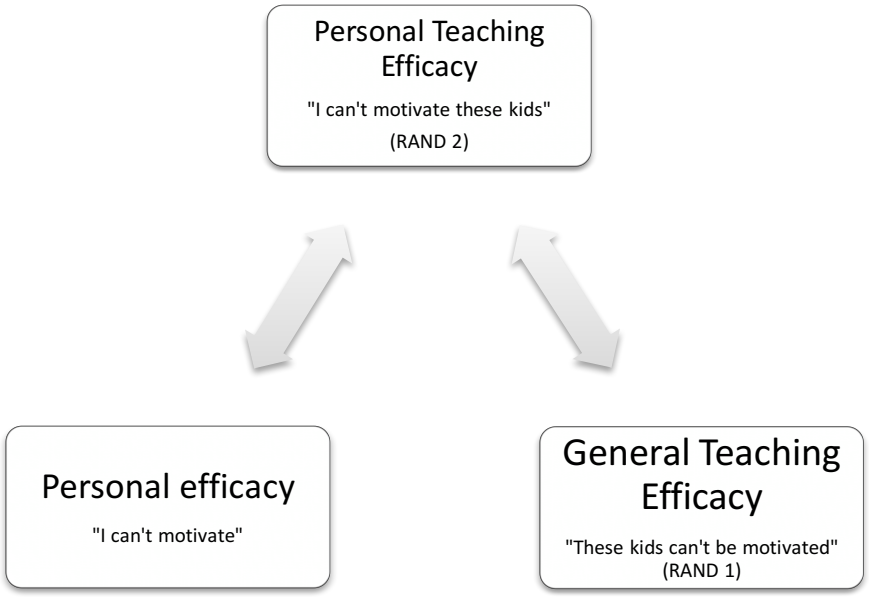

Figure 2. Ashton's model of teacher efficacy (1983)

Shown in Figure 2, Ashton identified three types of efficacy beliefs: teaching efficacy, personal efficacy and personal teaching efficacy. General teaching efficacy (GTE), conceptually related to Rand Item 1, represents teachers' views of the "general relationship between teaching and learning" (Ashton, 1983, p. 2), or beliefs that students can be motivated. Personal efficacy (PE), not represented in the RAND items, refers to teachers' belief that they themselves possess the capability and skills to motivate students[1]. Importantly, these two beliefs contribute to the third belief, their sense of personal teaching efficacy (PTE). PTE is teachers' beliefs that they can motivate students in their current

[1] In 1984, Ashton renamed this construct as "generalized beliefs about perceived self-efficacy" or "teachers' general sense of effectiveness as a person." Little discussion of this construct is provided; thus, I use the 1983 reference which provides more clarification as to how PE might relate to the RAND items

context. In other words, while PTE is the "best predictor of teacher behavior" (p. 2) because it is most context-specific belief, personal and general teaching efficacy are foundational beliefs that provide attributions for why teachers might have low PTE. Ashton implies, but does not state or empirically verify, that teachers with high PTE have both high personal and teaching efficacy.

Ashton and colleagues (1983) make an important point that identifying which of the two beliefs, PE or GTE, contribute to the origin of a teacher's sense of low PTE is critical because it determines the most appropriate intervention strategy to increase a teacher's sense of efficacy. A teacher who believes that students cannot be motivated will likely feel little responsibility to motivate those students (Lauermann & Karabenick, 2011) and will require different types of interventions than would a teacher who believes students can be motivated yet lacks the knowledge or skill to diagnose problems of student motivation or build relationships with students. This important point has been taken up by later theorists who have questioned how to translate teacher self-efficacy research into practice (Klassen et al., 2014), the meaning of teacher self-efficacy beliefs (Wheatley, 2005), and the notion that teacher responsibility is related to TSE but conceptually and empirically distinct (Lauermann & Karabenick, 2013).

However, Ashton and colleagues' empirical work did not support the conceptual framework Ashton proposed. Whereas it was proposed that low PE or low GTE might contribute to low PTE, Ashton operationalized teacher efficacy as a composite measure of the two RAND items, which assessed the constructs of GTE and PTE. Therefore, the concept of PE—a teacher's sense that s/he has the capability to motivate students in general—as a potential cause of PTE was not investigated before attempting to increase it

(Ashton, 1983), despite the authors' claims that knowing the source of low PTE was critical to knowing which strategies would be most successful in influencing teachers' sense of efficacy. Moreover, the concept of PE was abandoned in the subsequent and oft-cited publication (Ashton & Webb, 1986).

Thus, although Ashton and colleagues provided a conceptual model for the study of teacher efficacy, their conceptual framework was not supported by empirical evidence. Subsequent research focused on the development and validation of scales to assess teacher efficacy rather than developing or validating theoretical conceptualizations of teacher efficacy. Through this work, the idea of PTE as a contextualized belief grew stronger, yet the idea that teachers can have an effect on student performance began to be seen as a different construct, a divergence from how Armor and Ashton considered teacher efficacy.

Teacher efficacy and outcome expectancies

Through subsequent efforts to use factor analysis to identify dimensions of teacher efficacy, the idea began to take hold that teachers' beliefs about student performance and outcomes of teaching were separate from the construct of teacher efficacy. This shift resulted indirectly from efforts to clarify and reinterpret the notion of general teaching efficacy proposed by RAND and Ashton. One popular interpretation was that GTE was better described as outcome expectancy. According to social cognitive theory, outcome expectancies are the physical, social, or self-evaluative outcomes a teacher would expect given successful performance of a task.

Gibson (1983; Gibson & Dembo, 1984) first explicitly associated GTE with outcome expectancy[2]. Citing Bandura's 1978 work, Gibson (1983) argued that outcome

[2] Ashton and colleagues did not explicitly associate GTE with outcome expectancy, although later researchers credited them with this association (Guskey, 1994; Woolfolk & Hoy 1990).

expectancy would reflect the degree to which teachers believe good teaching can bring about positive changes in students given potential limiting background and contextual factors, an interpretation supported by other researchers (Soodak & Podell, 1996). This conceptualization was a distinct departure from RAND and Ashton, who conceptualized belief about student performance as an influence on teachers' personal efficacy beliefs. Furthering this conceptualization, Gibson and Dembo published the Teacher Efficacy Scale (TES; 1984), building off of items used in the RAND study. This scale is one of the most widely used scales to this day despite psychometric concerns of this scale (Klassen, Tze, Betts, & Gordon, 2011).

At the same time teacher efficacy researchers in science began to solidify use of the term "outcome expectancy" with the creation of the Science Teaching Efficacy Beliefs Instrument and, later, a mathematics version (Enochs, 1993; Enochs & Riggs, 1990; Enochs, Smith, & Huinker, 2000). These measures have two subscales of efficacy: personal teaching efficacy, measured by 13 items (e.g., *I will generally teach science ineffectively*) and outcome expectancy, the extent to which teachers believe that their teaching will lead to positive student outcomes (e.g., *The low mathematics achievement of some students cannot generally be blamed on their teachers*). These scales have been used extensively in science and mathematics teacher research (Andersen, Dragsted, Evans, & S\orensen, 2004; Bautista, 2011; Bursal, 2008; de Laat & Watters, 1995; Wu & Chang, 2006), likely contributing to the use of the term outcome expectancy.

I contend that these two interpretations of outcome expectancy as the same as GTE or as interpreted in the STEBI are inconsistent with Bandura's notion of an outcome expectancy for two reasons. First, an expectation for positive changes in students is not a

physical, social nor self-evaluative outcome experienced by the teacher. According to Bandura, an example of a teacher's outcome expectancy for successfully performing a teaching task would be decreased stress (physical), recognition from colleagues (social), or feelings of pride (self-evaluative). Second, outcome expectancies do not influence behavior because one can believe that positive outcomes will occur yet feel incapable of performing the task (Bandura, 1997). However, teachers with low beliefs that teaching can bring about student outcomes (as assessed by the STEBI) are likely to exhibit decreased effort and intention to leave the field (Edwards & Green, 1999; Malow-Iroff, O'Connor, & Bisland, 2004). This suggests that the belief that good teaching can bring about student outcomes is, in fact, a fundamental belief that influences teachers' sense of efficacy and subsequent behavior rather than an outcome expectancy, since outcome expectancies do not influence behavior.

The underlying implications of this conceptual confusion of what constitutes an outcome expectancy because of the reference to getting through to students was a significant change in the meaning of teacher self-efficacy beliefs from its original conceptualization: this new perspective proposes that the task is teaching and students are separate from that task. While outcome expectancy was the most common labeling of this other factor, other theorists referred to it as outcome efficacy and identified potential issues with its measurement.

Teacher efficacy and outcome efficacy

Woolfolk and Hoy (1990) suggested that the GTE factor was an efficacy expectation measured at the general level, not an outcome expectancy, because it involved

the potential of teachers to perform.[3] Instead, they proposed that outcome expectations could be "changes in student attitudes" or "recognition and rewards for the teachers involved" (p. 82). Again, I contend that only this latter proposition, recognition and rewards, is consistent with Bandura's notion of an outcome expectancy representing a social outcome that results from successful completion of a task.

Soodak and Podell (1996) used principal components factor analysis with varimax rotation on responses to a modified version of a teacher efficacy scale used by Woolfolk and Hoy (1990). Half of these items referred to positive outcomes and the other half to negative outcomes. Three factors emerged which the authors labeled as personal, teaching, and outcome efficacy. The authors defined personal efficacy as efficacy for enacting specific teaching behaviors such as adjusting an assignment or managing student behavior. Teaching efficacy referred specifically to the influence of the environment or heredity on student behaviors while outcome efficacy, consisting of four items, was interpreted as outcomes of effective teaching behaviors (e.g., *When a student gets a better grade than he usually gets, it is usually because I found better ways of teaching that student*). This interpretation of factors was supported by other researchers (Brouwers & Tomic, 2000).

In fact, all of the items from Soodak and Podell's "outcome efficacy" factor reference teacher attributions for student success to their specific teaching behaviors, or internal attributions for student success (Weiner, 1985). By contrast, items on another subscale mostly represented external attributions for failure to positively impact student learning. Guskey and Passaro (1994) provided more empirical evidence to support the internal/external dimension. Using the TES (Gibson & Dembo, 1984) plus four items from

[3] They would later reverse this stance in a 1998 publication with Tschannen-Moran and state that self-efficacy, by definition, cannot be measured at the general level of "all" or "most" teachers; it must be an individual belief.

Woolfolk & Hoy (1990) and the RAND items, principal components analysis was used to identify that the two constructs of PTE and GTE corresponded to an internal/external dimension (operationalized as can influence/cannot influence) rather than personal/teaching dimension. However, items from the internal scale are overwhelmingly positive attributions for success, while the items from the external scale are mostly negative attributions for failure. Thus, the corresponding positive and negative valences of the internal and external dimensions, respectively, act as a confound in interpreting these results.

In 2002 Brouwers and Tomic pointed out that all previous research had relied on use of principal components analysis, an exploratory procedure, which provided no information about the overall fit to a priori models of TSE. To address this gap, the authors used confirmatory factor analysis to test the fit of four specified factor models against null and saturated models with a randomly selected sample of 540 Dutch secondary teachers. The models comprised the proposed two-factor solution from Gibson and Dembo (1984), three-factor models proposed by Soodak and Podell (1996), and a four-factor model that added in efficacy for classroom management. While a four-factor model provided the best fit to the data, it did not meet the standard indices for adequacy, even after re-specifying the model and eliminating items to improve fit.

In sum, researchers generally converged on an understanding that the domain of personal teaching efficacy represented a contextualized belief about one's teaching skills. Disagreements surfaced in relation to the interpretation of what were considered other dimensions of teacher efficacy: attributions for student success and failure. While at first called general teaching efficacy, it was later interpreted as outcome expectancy and then

outcome efficacy. This shift in labeling beliefs related to student learning and performance as separate from personal efficacy beliefs marked a significant, yet not explicitly identified, shift in the meaning of teachers' self-efficacy beliefs away from the original conceptualization of teacher efficacy from RAND and Ashton.

Debate about interpretation of results of factor analytic studies and meaning of outcome expectancy and self-efficacy beliefs paved the way for Tschannen-Moran and colleagues to reconceptualize teacher self-efficacy and its measurement in two seminal articles published in 1998 and 2001. As a result of their comprehensive conceptual framework, the idea of outcome expectancy and general teaching efficacy as dimensions of teacher self-efficacy was dropped. This work is referred to as the second, or present, era of teacher self-efficacy research.

Teacher self-efficacy reconceptualized: The present era

It is important to describe the conceptualization of teacher self-efficacy from this second era of teacher self-efficacy research (1998 to present) and the measurement issues raised that are relevant for the present studies. Research conducted here: (a) is more closely aligned with a social cognitive theory of self-efficacy, yet (b) further away from measuring teachers' judgment of their perceived capability to influence student motivation and performance, and towards (c) conceiving TSE as teachers' perceived capability to enact teaching tasks that are important to create environments conducive to student learning. In this way, students are, at times, conceptually positioned as outcome expectancies of good teaching rather than part of the teaching task, though this may vary by domain of TSE considered. I shift here to using the acronym TSE to refer to teacher self-efficacy that is

strongly grounded in this theoretical tradition representative of the second era of TSE research.

Seeking to clarify the conceptual and measurement problems in the study of TSE, Tschannen-Moran and Woolfolk-Hoy (1998) comprehensively reviewed the literature and offered a refined definition of TSE that was rooted in social-cognitive theory: "Teacher efficacy is the teacher's belief in his or her capability to organize and execute the courses of action required to successfully accomplish a specific teaching task in a particular context" (p. 22). This conceptualization of TSE is distinct from the RAND conceptualization that proposes teachers are evaluating if they can positively influence student motivation and performance. In the definition proposed by Tschannen-Moran and colleagues, however, the task is less clear: is it teaching students, teaching strategies, or both? The source of ambiguity lies in the interpretation of "successfully." By what standards is success determined, and does that relate to positively influencing student achievement?

The conceptual framework for TSE proposed by Tschannen-Moran and colleagues somewhat clarifies how they consider teachers' judgments of what constitutes the teaching task and evidence of mastery of (i.e., success of) that teaching task. Their model is based a social cognitive perspective concerning the triadic and reciprocal relationship between the individual, the environment and their behavior (Bandura, 1977b). In this model, a portion of which is shown in Figure 3, the first two boxes from left to right are consistent with Bandura's model of self-efficacy: The four *sources of teachers' efficacy beliefs* and how they are interpreted, *analyses and interpretations*, clarify that it is how the teacher interprets these experiences (i.e., as successes or failures, or as helpful or not) and the

importance the teacher places on these experiences that makes them relevant sources of efficacy information (Bandura, 1997).

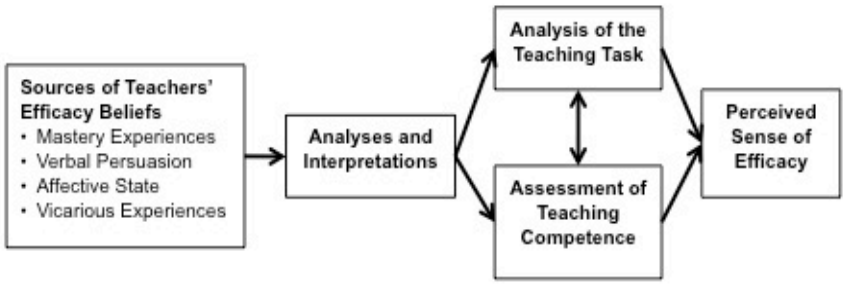

Figure 3. Portion of process model showing formation of teacher self-efficacy beliefs (Tschannen-Moran et al., 1998).

The two processes according to which teachers evaluate their perceived sense of efficacy for a teaching task are *assessment of teaching competence* and *analysis of the teaching task.* The *assessment of teaching competence* is the assessment of current functioning and is most similar to PTE. During this process, teachers judge their knowledge, skills and capabilities to meet the demands of the teaching task in their current teaching environment. This process does not assess past functioning, as some PTE items did, nor does it assess efficacy in hypothetical situations.

Although not specified in the model, teachers' judgments of their competence can be self-referenced or norm-referenced, in that teachers judge their competency based on internal standards for competence or external standards present in the environment. For example, Ashton, Buhr and Crocker (1984b) found in a sample of in-service teachers that their efficacy judgments as measured by RAND items were norm-referenced. Sixty-five

teachers enrolled in a graduate course responded to two sets of vignettes, one in which they were asked to rate how "effective" they were (self-referenced) and the other how effective they were compared to "other" or "most teachers" (norm-referenced). Only the norm-referenced ratings were significantly correlated with teacher self-efficacy as measured by the RAND items. The idea that TSE is norm-referenced has implications for what is considered a beneficial TSE outcome of professional development. Increasing the competency of teachers through professional development might actually make teachers feel less confident because they have higher norms for success available to them in the environment. Instead, most studies of professional development consider an increase in teachers' sense of efficacy as a desired outcome of the intervention (Abelson, 1997; Behnke, 2007; Bruce et al., 2010; Fahlman, Singleton, & Kliber, 2002).

The other process, the *analysis of the teaching task*, merits more careful consideration as the definition of what teachers and theorists consider to be the teaching task is the present research focus. The *analysis of the teaching task* is conceptualized as a means-end relationship in which the outcomes are instrumental in determining the actions required to accomplish them successfully (Skinner, 1996). This approach involves answers to two questions. First, as self-efficacy items are context-specific, goal-referenced evaluations of competence (Bong & Skaalvik, 2003), teachers ask themselves, "What outcomes do I seek—what is success in this teaching task?" (Tschannen-Moran et al., 1998, p. 232) In this way, the Tschannen-Moran model proposes that success is defined by the teacher for each specific teaching task.

Second, teachers consider, "What means or actions will be required to accomplish this particular teaching task—to succeed in this situation?" (Tschannen-Moran et al., 1998,

p. 232). The authors propose that the answers to these two questions are dependent on the teachers' perceptions of the resources and constraints available to them in their teaching context. Some of these resources and constraints include time available, access to materials, student abilities' and motivation, and managerial issues. Thus, the construct of GTE (teachers' beliefs about the ability of teachers to influence student achievement in the face of external obstacles) becomes subsumed within the teachers' analysis of the teaching task but altered in two important ways. First, teachers consider their personal capability to overcome these limiting factors, not just the general relationship between teaching and learning. Second, this judgment includes a consideration of not just constraints but also resources in a particular context. Resources in the environment may include community support, diversity and supportive leadership. Previous measures of GTE items typically assumed that the home environment could exert a negative influence on student performance and motivation.

In sum, according to this model teachers may perceive mastery experiences as those that (a) have a positive effect of student learning, or (b) are dependent on their own skill in enacting a teaching strategy independent of its effect on students. This potential source of variation is problematic not only because it implies different standards for the same task, but because it suggests that teachers may hold different conceptualizations of what their responsibility is as teachers (Lauermann & Karabenick, 2011; 2013), and thus, the teaching task against which they evaluate their sense of efficacy. An important consideration is whether some highly efficacious teachers feel that they are capable of performing certain tasks despite the fact that their performance fails to reach students in an

effective manner. Such unknowns pose problems in interpreting the meaning and practical implications of TSE (Klassen et al., 2014; Labone, 2004; Wheatley, 2005).

Teachers' conceptualizations of the teaching task and what constitutes evidence of success may also depend on the domain of teaching being considered. The following three sections consider three domains of TSE proposed by Tschannen-Moran and colleagues and supported by other researchers: classroom management, student motivation, and instructional strategies (Holzberger et al., 2013; Kunter & Baumert, 2006; Skaalvik & Skaalvik, 2007; Tschannen-Moran et al., 1998). After summarizing each domain of efficacy and how it has been conceptualized and measured, I conclude with whether teachers' judgments of success in the teaching task are likely to include evidence of success in achieving student outcomes.

Classroom Management Efficacy

Although Gibson and Dembo's Teacher Efficacy Scale (TES; 1984) contains items related to managing behavior in the classroom, Emmer and Hickman (1991) were among the first to distinguish classroom management and discipline efficacy as a conceptually and behaviorally distinct dimension from the "ability to influence learning or achievement outcomes" (p. 757). In particular, they identify that the goal of classroom management is to achieve order and cooperation, outcomes that are not immediately linked to student learning. Items developed were based on Doyle (1985) and his overview of organization and management of the learning environment.

The items comprising Emmer and Hickman's Classroom Management and Discipline Scale (CMDS), shown in Table 1, largely reflect the ideas of behavior management (e.g., Items 5, 11 and 14), structure and routines (e.g., Items 1, 3 and 15), and,

to a lesser extent, norms (Item 9). Unrepresented in these items is the co-construction of these norms between teachers and students as well as the demands of the content. Looking at factor fit, items representative of instructional strategies loaded onto the classroom management factor (e.g., Items 8, 10 and 13), but were then moved to the other factor. It was concluded that classroom management efficacy was a distinct dimension of teacher efficacy. Further psychometric work by Brouwers and Tomic (2001; Brouwers, Tomic, & Stijnen, 2002) confirmed this conclusion.

The classroom management efficacy subscale of the TSES (Tschannen-Moran & Hoy, 2001) is similar to the CMDS in that it emphasizes behavior management slightly more than it does structures and norms. In addition, neither scale emphasizes the dependent nature of the norms on the demands of the academic content. Half of the items focus on management of disruptive behavior (two of four items on short form; four of eight items on long form) and the other half on norms and routines. Unlike the CMDS, items assess what teachers "can" do to accomplish specific management tasks, rather than global assessments of classroom management skills (e.g., Item 4 from Table 1) or perceptions of knowledge (e.g., Items 11, 14, 16, and 17). Thus both the CMDS and the TSES, the most commonly used measures of classroom management self-efficacy, emphasize the management of behavior more so than the structures of organization, norms and the influence of academic content on those decisions (the forthcoming section "Structured Classroom Management" will expand on these ideas).

Table 1

Classroom Management and Discipline Scale (Emmer & Hickman, 1991)

Item number	Items	Factor loading
1	I know what routines are needed to keep routines running smoothly.	.65
2	I know what kinds of rewards are used to keep students involved.	.65
3	If students stop working in class, I can usually find a way to get them back on track.	.63
4	I have very effective classroom management skills.	.63
5	I can keep a few problem students from ruining an entire class.	.61
6	I can communicate to students that I am serious about getting appropriate behavior.	.61
7	I am confident in my ability to begin the year so that students will learn to behave well.	.61
8	If a student did not remember information I gave in a previous lesson, I would know how to increase his/her retention in the next lesson.[a]	.57
9	I find it easy to make my expectations clear to students.	.56
10	If one of my students couldn't do an assignment, I would be able to accurately assess whether it was at the correct level of difficulty.[a]	.53
11	If a student in my class becomes disruptive and noisy, I feel assured that I know some techniques to redirect him quickly.[b]	.50
12	When I really try I can get through to most difficult students.[a]	.48[c]
13	When a student is having trouble with an assignment, I am usually able to adjust it to his/her level.	.44[c]
14	There are very few students that I don't know how to handle.	.34
15	I don't always know how to keep track of several activities at once.	-.32[c]
16	Sometimes I am not sure what rules are appropriate for my students.	-.32[c]

[a] Item from Gibson and Dembo PTE, moved to scale 3 (PTE)
[b] Item from Gibson and Dembo PTE, retained on this factor
[c] Cross loadings above .3 with one or two more scales

Finally, the items in the CMDS in Table 1 clarify how teachers' judgment of their classroom management efficacy may likely include evidence of success with students. Item 5 directly assesses confidence a teacher can achieve a result related to student behavior (keep disruptive students from ruining a class) rather than a strategy that might achieve that outcome (e.g., keep pace of class moving appropriately, arrange seating to maximize on-task behavior between students, etc.). Therefore, teachers' judgments of their efficacy for this task may likely include evidence of having achieved that outcome with students (and potentially whether it is possible to achieve such outcomes with students).

Student engagement and motivation self-efficacy

Teachers' beliefs that they are capable of motivating their students may be one of the main pathways by which they influence students' academic and cognitive development (Bandura, 1997). As stated earlier, Ashton was the first to identify teachers' sense of efficacy as an implicit personal theory of student motivation (1983), yet this conceptualization of TSE is a generalized belief about locus of control rather than a task-specific view of self-efficacy to motivate and engage students. Teacher self-efficacy for motivation is typically conceptualized and measured as motivating individual students to value learning or believe they can do well in a particular class (e.g., Skaalvik & Skaalvik, 2007; Tschannen-Moran & Hoy, 2001).

The development of the TSES was the first attempt by researchers to capture the construct of student motivation and engagement. Items for this subscale (and others) were developed through discussion among researchers and teachers about important tasks for teachers. Bandura's unpublished teacher efficacy scale (n.d.) also provided a foundation from which to further develop items. No specific theoretical framework related to student

motivation or engagement was used. Items developed included general strategies such as motivating students to do well, fostering creativity, and helping students think critically.

Blazevski (2006) was the first to draw on existing theory of student motivation to develop a scale for teacher efficacy for student motivation. Blazevski noted that some items from the TSES subscale for student motivation are not actually specific to student motivation (e.g., fostering creativity, getting through to difficult students). Through reviewing other developed instruments (MAST, TSES, RAND, and STEBI), she developed a 6-item scale specific to fostering student motivation in mathematics (e.g., get students excited about math, getting through to unmotivated students). Using this scale, middle school teachers' self-efficacy for supporting student motivation was indirectly related to student achievement in mathematics. It was also a significant predictor of variance in student self-efficacy for mathematics and students interest in mathematics. Additionally, years of experience was a significant negative predictor of TSE for student motivation (Blazevski, 2006). This scale was unpublished, and thus has not been used in further studies.

Also drawing upon the TSES, Hardré and Sullivan (2008, 2009) developed the Motivating Strategies Questionnaire (MSQ) comprised of two subscales: efficacy for diagnosing motivating needs and efficacy for responding appropriately to this need. Efficacy for diagnosing needs emerged as a strong contributor to the majority of strategies teachers employed to motivate students, such as providing emotional support, appealing to relevance and value, connecting content to aspirations and future goals, and acknowledging peer pressure. Follow-up interviews showed that teachers lacked knowledge about how to motivate students, yet felt more confident to diagnose rather than

treat motivational problems. Duffin (2010) found that pre-service teachers have a lot of declarative knowledge about how to motivate students, yet tended to rely on task-extrinsic rewards. Overreliance on task-extrinsic rewards may undermine intrinsic motivation (Lepper, Corpus, & Iyengar, 2005; R. M. Ryan & Deci, 2000), suggesting that teacher may indeed lack knowledge about how to effectively motivate students. Declines in TSE for motivating students as teachers gain experience (Blazevski, 2006) may be related to teachers' realization through experience that they lack this knowledge.

In sum, most existing measures of TSE for motivating and engaging students are contextualized versions of the TSES. Items from the TSES may not clearly reference working with students who exhibit motivational problems (e.g., "difficult students"), and items from the long version of the scale include items that may be more relevant to instruction (e.g., fostering creativity and critical thinking).

Finally, similar to the domain of classroom management efficacy, items typically ask about achieving a desired behavior or attitude from students (e.g., motivating a student, helping them to believe, etc.) rather than using particular strategies that might lead to success in engaging and motivating students, (e.g., making content relevant to students' lives, forming personal relationships with students, etc.) Thus, teachers may more often use evidence of student success to judge their self-efficacy because of how items are phrased.

Instructional self-efficacy

Instructional self-efficacy is the least conceptualized domain of TSE. Few scales specific to studying this domain of TSE have been developed with the exception of the instructional strategies subscale of the TSES (Tschannen-Moran & Hoy, 2001) and the

personal teaching efficacy scale of the Science Teaching Efficacy Beliefs Scale (STEBI; Enochs & Riggs, 1990). Thus, this section will analyze the items on these subscales to infer what theoretical frameworks or influences are reflected in the assessment of instructional self-efficacy.

The TSES was developed in conjunction with current and past teachers and researchers who developed items that "represented important tasks or elements of teaching" (p. 796). Additionally, Bandura's unpublished and un-validated Teacher Self-efficacy Scale (n.d.) was used as a basis for developing items. Analysis of the items on the instructional strategies subscale shows that this scale assesses teachers' self-efficacy to plan for instruction and work with students both in and out of the classroom, considered to represent pre-active and interactive components of teaching (Jackson, 1990). Most tasks are those that can be done in class (interactive) or in preparation for class (preactive). Specifically, TSES items assess teachers' perceived capabilities to respond to difficult questions from students, develop appropriate challenges for capable students, gauge student comprehension, use a variety of assessment strategies and craft good questions from students. These tasks can be preactive or interactive because, for example, crafting questions or adjusting lessons can be done in the moment as the need arises or before class as the teacher prepares lessons. As teacher preparation programs tend to focus on preactive components of teaching (Grossman, Hammerness, & McDonald, 2009), and there is less time pressure and cognitive demand on teachers as they plan for instruction compared to engaging in instruction, teachers may feel more efficacious if they consider only the preactive side of tasks. It is unclear whether teachers consider both elements for all tasks or

what other factors might influence their interpretation of TSES items, which is the subject of the empirical investigation in Chapter 4.

Moreover, items from the TSES do not make explicit reference to managing demands of the content, considered a component of effective instruction in the classroom (Newmann, 1996). As proposed by Cohen, Raudenbush and Ball (2003) teaching is interactive, defined as "what teachers do, say, and think with learners, concerning content, in particular organizations and other environments, in time" (2003, p. 124). In this framework, shown in Figure 4, the teacher, the students and the content constantly interact and influence each other. Within this framework for teaching, one can never think about teaching without the content of what is being taught (D. K. Cohen, 2010). While much of what teachers do with instruction is determined by the content being taught, it is unclear whether teachers consider their efficacy for content being taught (e.g., knowledge and practices) when thinking about their capabilities to instruct. This question frames the qualitative investigation in Chapter 4.

The Science Teaching Efficacy Beliefs Instrument (STEBI; Enochs & Riggs, 1990) closely assesses teacher self-efficacy for instruction, and as the name implies, it is meant to be content-specific. Yet few instructional tasks from this scale can be considered unique to the knowledge and practices of that domain. For example, a number of items from the personal teaching efficacy subscale of the STEBI ask teachers to rate their general belief that they can teach science effectively. Other items get more specific, yet are tasks that teachers in any subject would need to do, such as answering students' science questions or welcoming student questions. Only two items ask about efficacy for tasks unique to

teaching science: monitoring science experiments and explaining to students why science experiments work.

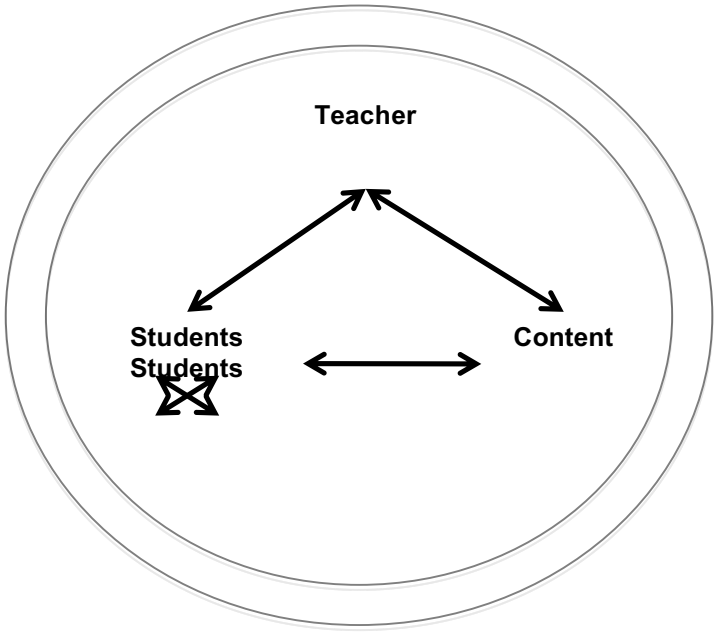

Figure 4. Conceptual Model for Teaching as Interaction (Cohen, Raudenbush & Ball, 2003)

Moreover, while no particular theoretical frame is assigned to the STEBI, it is evident that few items reflect a constructivist or inquiry-based view of science instruction. For example, no items ask about facilitating student construction of their own explanations about why science experiments work, while one STEBI item asks about teachers' efficacy to explain why science experiments work. Importantly, teachers must explain content even in inquiry-based classrooms (Hmelo-Silver, Duncan, & Chinn, 2007); my point here is that the STEBI does not assess other important aspects of inquiry-based instruction such as facilitating students' abilities to construct scientific explanations. Supporting this

interpretation of the STEBI, teacher efficacy for inquiry-based instruction has been investigated as a separate construct (Cripe, 2009; Nie, Tan, Liau, Lau, & Chua, 2013).

The TSES appears to have neither a didactic nor constructivist view of instruction. In fact, items may be interpreted in either paradigm. For example, a teacher's view of what good questions are may vary between questions that are designed to inspire higher-level thinking to questions that elicit students' thinking about the content, whether or not it was at a higher-level or not. Chapter 4 investigates teachers' responses to more vague items such as these. Additionally, it is unclear how much teachers' knowledge of the content and students' thinking about the content factors into their efficacy judgments. Such potential sources of variation may be problematic for the validity of the instructional efficacy subscale of the TSES; in fact, a recent study found that this subscale was more highly correlated with general-self efficacy than the other dimensions of TSES, suggesting that the instructional strategies subscale might be too general to hold predictive value (Pfitzner-Eden, Thiel, & Horsley, 2014). Relatedly, teachers' instructional efficacy is typically the highest rated domain of efficacy compared to efficacy for student engagement and classroom management and shows less variation in scores (Chong, Klassen, Huan, Wong, & Kates, 2010; Klassen & Chiu, 2010; Knoblauch & Chase, 2015), though some studies show that classroom management efficacy is higher (Fives & Buehl, 2009).

Finally, compared to other domains of TSES, teachers may be less likely to reference student outcomes (e.g., achievement or learning as an indicator of success). One reason may be that, as depicted in the Instructional Triangle in Figure 4, teachers have to consider content in addition to students when considering instruction. Perhaps some teachers focus more on their efficacy to teach content and less on students and their

learning outcomes. Furthermore, instructional efficacy items as represented by contemporary scales like the TSES and the STEBI do not typically assess teacher efficacy for outcomes such as helping students to learn or retain material. Note that this is in contrast to earlier conceptualizations of teacher efficacy in which helping students learn and achieve was the task (e.g., RAND, TES).

In summary, of the three TSE domains—classroom management, student engagement and instruction—the latter domain is the least well conceptualized. Classroom management efficacy has a rich tradition of study in educational psychology, while instructional self-efficacy is the least conceptualized and studied. These three domains of TSE are similar, yet distinct, to three domains of teaching that have recently been identified as core domains of teaching (Klieme, 2012): structured classroom management, supportive classroom climate, and cognitive activation. However, it may be that the tasks commonly assessed in teacher self-efficacy measures do not sufficiently capture the most critical, research-based views about what effective teachers do in the classroom (O'Neill & Stephenson, 2011). General teacher self-efficacy measures, such as the TSES, must be careful to strike the right balance between specificity and generality, a common debate in the larger field of self-efficacy research (Bong, 2006; Schunk & Pajares, 2009), yet they must also capture the most critical tasks that are important to high quality and rigorous instruction to hold predictive value. Understanding to what degree there is alignment between teachers' domain-specific sense of efficacy and domain-specific teaching behaviors known to be effective for student learning is a crucial component of understanding the meaning and implications of TSE.

Domains of Effective Instruction

The following three sections delve deeper into the three domains identified as important for effective classroom teaching and their conceptual congruence with teachers' sense of efficacy for those domains. I focus on studies in which effective teaching has been defined as teaching strategies that are associated with student learning gains, drawing from educational and psychological literature. Each section is structured in three parts: 1) a review of the educational research on the teaching domain, 2) an analysis of the degree of congruence between the corresponding domain of teacher-self efficacy, and 3) a summary of empirical relationships between this domain of effective teaching and teachers' sense of efficacy. The overviews in parts 1 and 2 provide context for a review of the empirical associations between both teaching practices, part 3.

Effective teaching domain one: Structured classroom management

The first domain of effective classroom teaching shown in Table 2, structured classroom management, describes a classroom environment that runs "like a machine" because students and teachers are moving efficiently towards specific goals set by the teacher and communicated effectively to students. This domain of effective teaching was one of the first to be studied in connection with student achievement (Evertson et al., 1980). Though typically interpreted as behavior management, in actuality teachers who score highly in this domain were not only observed to effectively and efficiently deal with student behavioral issues, they also proactively managed the environment through setting up routines and procedures, cuing students as to situational expectations and consistently enforcing rules (Brophy & Good, 1986). Moreover, these teachers are effective communicators, delivering directions and explanations clearly and effectively. Also

represented in this domain is the clarity and organization of instruction: to what extent lessons, learning goals, and assessments work together to present a clear and coherent set of learning activities for students. Curricular alignment is strongly associated with student achievement (S. A. Cohen, Hyman, Ashcroft, & Loveless, 1989).

Doyle (1985) also emphasized that classroom management is fundamentally a process of solving the problem of order in the classroom, not responding to individual behavior, problems of misbehavior or student engagement. He also points out that it is dynamic, a "harmony of action with structure and purpose" (p. 424). This "program of action" is jointly enacted by students and teachers and depends on norms for social participation and demands of the academic content (Doyle, 1985). In summary, Doyle and other theorists envisioned classroom management as a complex and highly structured system co-created by students and teachers that is governed by norms for interaction and behavior and influenced by the demands of the content.

Comparing structured classroom management with how TSE for classroom management has been conceptualized, this domain represents a broader interpretation of the tasks that contribute to managing a classroom than simply managing behavior. It is also assumed that teachers must be able to manage the space, content and materials in the classroom jointly with students. Classroom management self-efficacy typically reflects teachers' capabilities to manage disruptive behavior, get students back on task working and, to some extent, set up routines and procedures (O'Neill & Stephenson, 2011). A final difference lies in the extent to which content is presented in a structured logical way that is comprehensible to students, a factor that is absent from the conceptualization of TSE for classroom management.

Allinder (1994) found that high efficacy teachers demonstrated greater planning and organization, an important component of structured classroom management. Among studies that have used validated observational measures, findings vary and may depend on the focus of the observational tool. Using Danielson' Framework for Teaching observational measure (Danielson, 1996), Heneman and colleagues (2006) found that classroom management efficacy (CME) was associated with coherent and logical design of teaching (r=.33), a domain of planning and preparation. However, using the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008) with 509 teachers, Jamil (2012) found no association between CME and classroom organization, a domain encompassing behavior management, productivity and instructional learning formats.

Table 2

Structured Classroom Management indicators synthesized from literature

Structured Classroom Management
A. Opportunity to learn (classroom organization)
+ momentum, brisk pacing of lessons
+ effective transitions
+ routines and procedures for learning
+ high student engagement (behavioral)
B. Proactive behavior management
+ nipping potential problems in the bud
+ consistency of rule enforcement
+ providing clues to situational expectations
+ effective treatment of interruptions
C. Instructional coherence
+ clear instructional goals
+ curriculum alignment
+ coherent content (coherent explanations, clear directions, vivid descriptions)
+ goal-oriented assessments
+ ongoing formative assessment

Focusing specifically on studies assessing classroom management efficacy as a distinct dimension of teacher self-efficacy, Emmer and Hickman (1991) found that CME was associated with preference for the use of positive strategies in the classroom such as adapting assignments to students' needs, praise, encouraging effort and having students develop plans for change. Other studies have found that CME is associated with a less bureaucratic orientation towards pupil control (Woolfolk & Hoy, 1990).

In relation to observed classroom practices, research using study-specific observational measures have found moderate associations between classroom management efficacy and use of generic, but not specific praise of in-service teachers (Reinke, Herman, & Stormont, 2013). However, Emmer and Hickman (1991) found no association between supervisors' assessments of pre-service teachers' use of positive strategies ($r = -.06$). Similarly, Rohs (2007) found no significant correlation between CME and the Classroom Practices Inventory, a measure designed to assess child-centered classrooms with positive emotional climates in elementary schools. In all some evidence exists that CME may be associated with teachers' self-reported and observed use of praise in the classroom, a strategy, which may speak more to the emotional climate in the classroom rather than the management of learning structures.

Table 3

Supportive Classroom Climate indicators synthesized from literature

Supportive Classroom Climate
A. Teacher-student interaction
+ teacher enthusiasm, warmth (e.g. smiling)
+ praise
+ teacher knows students as individuals and as learners
+ responsiveness to student input (e.g., integration of student ideas)
+ appropriate expectations
+ orientation to students' personal needs
+ process emphasized as much as product (mastery vs. performance goals)
+ responsive to individual differences: seen as opportunities for learning rather than obstacles to be overcome
+ may teach self-regulation strategies, if needed
+ student non-cognitive outcomes of self-efficacy, agency, and self-monitoring
B. Student-Student interaction
+ respectful interactions
+ mutual respect
+ cooperation
+ high student engagement (affective)
C. Structure of activities
+ cooperative learning
+ heterogeneous grouping
+ flexible grouping arrangements
+ adaptation of represented materials to the characteristics of the context/students

Effective teaching domain two: Supportive classroom climate

A supportive classroom climate (Table 3) speaks to the extent to which the teaching is in tune with student needs and creates a safe environment for learning. Three areas of focus are positive relationships between teacher and students, positive relationships between students, and heterogeneous structures for learning.

The degree of closeness or warmth in relationships between teachers and students in the classroom has long been studied, and beyond achievement has also been positively associated with student self-concept, self-efficacy, engagement and student expectations for success (Dolan & McCaslin, 2008; Eccles & Roeser, 1998; Wentzel, 2002). Early studies focused mainly on teacher actions that conveyed a sense of warmth, such as frequent smiling (Kounin, 1970). enthusiasm, and use of praise rather than criticism (Evertson et al., 1980). Extending this research, the appropriateness of expectations teachers have for children has been shown to be important (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996); moreover, to what degree they are autonomy supportive, or oriented towards students' personal needs and individual differences and willing to integrate student ideas into the classroom (Pianta, La Paro, Payne, Cox, & Bradley, 2002). Also important is that teachers view individual differences as opportunities for learning rather than obstacles to be overcome (Corno, 2008), and that there are student-to-student interactions in which students interact with each other respectfully, cooperate, and see each other as resources for learning (Pianta & Hamre, 2009; Rohrkemper & Corno, 1988). Moreover, by structuring activities that allow for cooperative work and teaching children how to work together productively, teachers have a great influence on the emotional

climate in the classroom and the creation of a supportive community of learners (Solomon, Battistich, Watson, Schaps, & Lewis, 2000).

In comparison to TSE for motivating students, the domain of supportive classroom climate focuses more broadly on creating an environment for learning among students. Student to student interactions and the pedagogical strategy of building relationships with students (Garcia & Shaughnessy, 2015) is a key component of this domain that is missing from TSE for student engagement. However, it must be noted that TSE items often assess the outcome of motivating students and not the means (e.g., pedagogical strategies) by which teachers might achieve that outcome. (Note: more is said about this in the following section, *Measuring multiple domains of teacher self-efficacy* as well as a focus of Chapter 4**)**. Teachers may consider the pedagogical strategies of creating a supportive classroom climate as they evaluate their self-efficacy for the outcome of motivating students. Thus, it is unclear to what extent there is a disconnect in the conceptualization of TSE for motivating students and the parallel domain of effective classroom teaching, supportive classroom climate.

Early studies of TSE using the TES scale (Gibson & Dembo, 1984) found that higher efficacy teachers were more likely to demonstrate some aspects of a supportive classroom climate. Teachers were observed to persist with students who were struggling (Ashton & Webb, 1984), set higher expectations for students (Coladarci & Breton, 1997), and be open to innovations that might better support student needs (Allinder, 1994; Cousins & Walker, 2000). More recent studies using the CLASS observational tool (Pianta & Hamre, 2009) found a positive relationship between teachers' self-efficacy and the emotional climate of the classroom (Guo, Connor, Yang, Roehrig, & Morrison, 2012; Guo,

Piasta, Justice, & Kaderavek, 2010), though other studies have found no relationship (Collmann, 2012). Using two-level cross-logged structural equation modeling, Holzberger (2013) found that teachers' self-efficacy ratings predicted their own ratings of individualized learning support one year later, but not their level of cognitive activation or classroom management in the classroom. TSE beliefs, however, did not predict nor were they impacted by student ratings of individualized learning support.

Effective teaching domain three: Cognitive Activation

Cognitive activation, shown in Table 4, is an area of effective teaching that was largely overlooked until late seventies and eighties and has been conceptualized in different ways in Table 4educational psychology and education research. Educational psychologist typically use the phase cognitive demand to mean the types of thinking students are asked to do (e.g., higher level or lower level thinking; Krathwohl, 2002) which can be assessed through teacher questions, activities and assessments (Boston & Smith, 2009; Holzberger et al., 2013). The role of thoughtful discourse around content has been highlighted through attention to the opportunities teachers give for students to answer questions with multiple answers or to answer questions that demand higher level thinking, particularly in the area of mathematics instruction (Cobb, Boufi, McClain, & Whitenack, 1997; Kazemi, 1998).

Furthermore, instruction in metacognitive strategies, such as reflecting on one's thinking and determining best strategies by which to solve problems, has been shown to strongly relate to student achievement (Williams et al., 2002), particularly the extent to which teachers emphasize the processes used to generate more that just the correct answers (Darling-Hammond & Bransford, 2005). This has often been investigated through having

students solve problems that are appropriately challenging (Kazemi, 1998; Newmann, 1996).

Some theorists have pointed out that the content being taught plays a critical role when considering the domain of cognitive activation. While higher-level thinking, substantive discussion, and project-based work, and content-based interactions are important, these interactions should also reflect legitimacy, accuracy and authenticity (Newmann, 1996). In other words, it is not only that students are asked to synthesize material and construct their own understandings, but that those understandings reflect central tenets of the discipline under study. For example, a teacher holding a discussion on apartheid in an 11th grade Social Studies would demonstrate legitimacy, yet this same discussion held in a mathematics classroom would not be legitimate. Similarly, a conversation about South Africa that demonstrated higher-level thinking but largely reflected incorrect information and connections to the real world would not be accurate. The ideas of legitimacy, accuracy and authenticity have been taken up by theorists in both domain-general (Newmann, 1996) and domain-specific (Ball, 2000; Kazemi, 1998) pedagogical theory.

Table 4

Cognitive Demand indicators synthesized from literature

Cognitive Demand
A. Thoughtful discourse
+ higher level cognitive questions
+ teachers probe for explanations
+ constructive feedback
+ mastery assessed through extended responses (orally or in writing)
+ public questions (response opportunities)
+ pose questions with multiple answers
+ students discuss ideas by talking (construction of knowledge)
+ encouraging students to relate classwork to their own experiences
+ careful attention and diagnosis of students' knowledge
+ scaffolding students' ideas and task involvement
+ high student engagement (cognitive)
B. Metacognitive strategies
+ reflection
+ prompt students to explain thinking and how they arrived at conclusions
+ strategy teaching (models and instructs learning, problem solving)
+ debriefing learning activities
+ establishing learning orientations (clarifying intended outcomes and cuing desired learning strategies)
C. Problem solving
+ Inquiry models & authentic activities
+ practice/application
+ interpretation of findings
+ opportunities for extended writing projects

As stated earlier, Cohen, Raudenbush & Ball (2003) propose that teaching is interaction among student, teachers, and content. Within this framework for teaching, one can never think about teaching without the content of what is being taught (D. K. Cohen, 2010). In other words, content should not be relegated only to the domain of "cognitive activation" but also to structured classroom management and student engagement. Of the two domains of Structured Management and Supportive Classroom Climate, only the domain of Structured Management contains a reference to the content being taught. *Instructional Coherence* is a subdomain of effective classroom teaching that refers to the clarity of the content and the alignment between assessment and goals. While this does not reflect the ideas of legitimacy and authenticity (Newmann, 1996), it does acknowledge that content is also a resource which teachers can use to effectively structure and manage a classroom. However, as noted earlier the assessment of TSE for classroom management rarely assesses teachers' sense of efficacy to structure the content. Therefore, there may be weaker demonstrated correlation between these domains, a proposition explored in Chapter 3.

Cognitive activation and TSE for instructional strategies are similar in that they focus on creating challenges for students that will prompt them to think at higher levels. TSE for instruction, however, does not specifically assess teachers' self-efficacy to inspire thoughtful discourse among students or use metacognitive strategies. However, as in the case of TSE for student engagement, teachers may think of these components as they evaluate their self-efficacy to "craft good questions" or "create challenges for capable students" (Tschannen-Moran & Hoy, 2001).

The connections between TSE and cognitive activation in the literature are inconsistent and generally weaker compared to other domains of effective classroom teaching. Studies using a 4-item German TSE scale (Schmitz & Schwarzer, 2002) have shown a positive correlation between TSE and cognitive activation, even over time and teacher versus student ratings (Holzberger et al., 2013; Kunter & Baumert, 2006); yet these correlations were weakest for cognitive activation compared to classroom climate and classroom management. Studies using the TSES, however, found no correlations between instructional self-efficacy and observed teacher instructional practices as assessed by the CLASS (A. M. Ryan, Kuusinen, & Bedoya-Skoog, 2015; Yan & Cheng, 2015). Just as the demands of a particular content area, such as science, might influence the assessment of TSE in that field, so might the unique demands of each domain influence how it is measured.

While it is important to assess the specificity matching between domains of TSE and effective teaching, it is also important to consider how best to assess teacher self-efficacy. The following section describes the current view of the multi-dimensionality of TSE and its implications for the measurement and meaning of TSE.

Measuring multiple domains of teacher self-efficacy

Having previously noted that popular scales like the TES were constructed to assess a confusing mix of past and present functioning as well as beliefs about knowledge rather than capabilities, Tschannen-Moran and colleagues (2001; 1998) modifying an unpublished TSE scale (Bandura, n.d.) constructed future-oriented, task-specific TSE survey items. All items contain the word “can” in reference to a specific action-oriented skill such as calming a student, crafting good questions, or helping students value learning.

Only one item is reminiscent of the previous era of teacher self-efficacy, and only appears on the long form (24 items) version of this survey: "How much can you do to get through to the most difficult students?" Although Bandura proposed this as an item for a subscale of instructional self-efficacy, the item is included in the student engagement subscale of the TSES, although with the lowest factor loading on this subscale (.47).

The focus on developing items more conceptually aligned with social cognitive theory and best recommendations for the assessment of self-efficacy beliefs was undoubtedly an advance for the field of TSE research. Lack of attention, however, was given to the nature of teaching and what constitutes a task in this professional domain, particularly in relation to students. One can see, therefore, differences in conceptualization of the task inherent in the subscales of the TSES, particularly in the short, 12-item version of the scale.

Items in the classroom management and student engagement subscales of the TSES ask teachers to what extent they can achieve a specific behavioral or attitudinal outcome (e.g., *How much can you do to control disruptive behavior in the classroom?*). The strategy by which they would achieve such an outcome is left unclear. In these items, therefore, the teaching task is construed as achieving some sort of desired change in students. These items are more conceptually similar to that of the RAND items just at a finer level of specificity: "getting through" to students is defined as controlling their disruptive behavior or helping them to value learning. I refer to these items as student outcome-based efficacy items.

In contrast, items from the instructional strategies efficacy (ISE) subscale ask teachers to evaluate their capabilities to enact a strategy (e.g., *To what extent can you*

implement alternative strategies in your classroom?) rather than capability to use a strategy to attain a student outcome (e.g., *To what extent can you successfully implement alternative strategies to increase student understanding?*), or simply to attain an outcome (e.g., *To what extent can you increase student understanding?*). With their lack of reference to student outcomes, I refer to these items as strategy-based items.

It is unclear if this difference in strategy versus outcome-based items on the TSES was intentional on the part of the authors to reflect the nature of the instructional domain and the introduction of the content that is to be taught. Moreover, it is unclear whether teachers' efficacy judgments in these domains mean different things and differentially predict their behavior. In other words, teachers' conceptualization of the task and evidence of mastery for classroom management may be quite different than how teachers construe the task and evidence of mastery for content-related instruction. For example, the task of classroom management may be considered to get students to comply with rules: evidence of mastery would be clearly linked to evidence of student compliance thus construing the teaching task as teaching/getting through to students. On the other hand, the same teacher might consider the task of instruction as giving clear and logical directions, and evidence of mastery might be based on personal or professional criteria for competent performance. In Chapter 4, I carefully compare teachers' conceptualizations of the teaching task and evidence of mastery across different domains of teaching tasks, focusing closely on instruction.

Despite these potential issues with the conceptualization of the teaching task in the TSES, it is clear that its development was a boon for TSE research in this second, present era, just as the TES (Gibson & Dembo, 1984) was a boon for the first era of teacher

efficacy research. The TSES is the most widely used measure of teacher self-efficacy and demonstrates strong psychometric properties, even across cultural contexts (Klassen et al., 2009, 2011). The development of short and long versions of the scale, and its easy accessibility likely contributed to it becoming one of the most widely used TSE surveys (Klassen et al., 2014, 2011). Other measures (e.g., Blazevski, 2006) have been developed based on modifications to the TSES because of its strong conceptual underpinnings in social cognitive theory. While other measures of TSE used in this era (e.g., Holzberger et al., 2013; Skaalvik & Skaalvik, 2007) are similar to the TSES in that they are future-oriented, assess perception of individual capabilities, and assess efficacy in multiple dimensions of teaching tasks.

While I have proposed that the nature of the task might be different among different domains of teaching task, it is also true that how teachers perceive the task and evidence of success might vary based on teachers' development of professional knowledge, that is, their understanding of task purpose and complexity and ability to perform them in the classroom. The following section reviews the role of teacher experience and school context on efficacy judgments.

The role of teachers' experience and context

Researchers have suggested that self-reported efficacy may mean different things and lead to different understandings of survey items for novice and experienced teachers (Henson, 2001; Tschannen-Moran et al., 1998; Wheatley, 2005). From a cognitive perspective, novice teachers' thinking is typically characterized as focused on retrieval of pedagogical strategies that have yet to become automatized. In complex and cognitively

demanding situations such as teaching, novice teachers may be more likely to focus on their own behaviors (How should I teach in this situation?) rather than on actual student outcomes (How are students responding to my teaching?), as a means of reducing the complexity of the classroom (Feldon, 2007a). In such an evaluation, their perceived level of pedagogical knowledge—knowledge of teaching strategies—might be a more salient factor in their efficacy judgments. In other words, capability to generate a lengthy inventory of strategies may lead to a higher sense of efficacy even if those strategies have not yet been successful in bringing about desired student outcomes. On the contrary, other scholars have suggested that novices may base their efficacy judgments on what they believe is possible rather than what they know how to do (Tschannen-Moran et al., 1998). It is unclear, and a goal of the present study to investigate, the extent to which these patterns of thinking are present and consistent across and within novice teachers, different domains of teaching tasks (e.g., instruction, student motivation, and classroom management), and different construction of the self-efficacy indicator (i.e., outcome or strategy-based indicators).

Another reason for this lack of novices' attention to student learning outcomes is that it depends on the skill to elicit student thinking and interpret such thinking with accuracy. Such skills develop over time but can also be taught to novices with explicit support and teaching (Ball & Forzani, 2010; Coffey, Hammer, & Levin, 2009; Thompson, Windschitl, & Braaten, 2013). Gabriele and Joram (2007) found in a think-aloud study that, relative to more experienced teachers, inexperienced elementary teachers focused less frequently on evidence of students' thinking to judge their success. With support from teachers, Mulholland and Wallace (2001) documented one novice's shift to focusing on

student learning outcomes. While the authors attributed this change to changing expectations for student on-task behavior, it is also possible that the support of the more experienced teachers around her facilitated her skill at interpreting student thinking. Importantly, these studies were conducted in the context of reform-oriented mathematics classrooms that provide explicit support for developing student thinking. To understand the generalizability of novice and more experienced teachers' thinking about student outcomes and their efficacy, it is imperative to study teacher cognition across different academic disciplines and in contexts in which they are not coached in a particular pedagogy.

Moreover, experienced and knowledgeable teachers may judge their sense of efficacy in distinct ways from their less experienced colleagues. Expertise is characterized by an ability to recognize patterns that aids in efficient problem-solving (Berliner, 2004; Feldon, 2007b; Sabers, Cushing, & Berliner, 1991). Expert teachers, therefore, are knowledgeable about patterns in student behavior, engagement and thinking and can use that knowledge to prevent problems (e.g. disruptive behavior) from occurring in the first place or improvise as problems arise (Borko & Livingston, 1989). As a result, they likely have a greater perception of how teacher actions affect student actions, which may lead to more controllable attributions for failure and success compared to novices (Mavropoulou & Padeliadu, 2002). As applied to instruction, expert teachers can anticipate patterns in student thinking about content, with particular attention to student misconceptions (Lampert & Clark, 1990). Thus, expert teachers may hold different, more complex standards for "success" in teaching tasks (Duckworth & Yeager, 2015) and also demonstrate more controllable attributions for success and failure with students, both

possibilities which warrant investigation into whether low efficacy for a novice teacher means the same thing as low efficacy for an experienced teacher.

Conclusion and looking ahead

In all, although the measurement and conceptualization of TSE has been refined to solidly reflect the theoretical bases of self-efficacy theory, there is insufficient research regarding the meaning of teachers' self-efficacy beliefs in relation to teachers' beliefs about student performance and teachers' use of research-based strategies known to be effective for student learning. Moreover, ambiguity remains regarding whether current measures of TSE capture research-based strategies of instruction. Finally, it is crucial to understand how these beliefs might differ not only in strength but in meaning based on teachers' prior experience teaching and context in order to understand the implications of teaching self-efficacy beliefs and how best to measure these beliefs.

Having reviewed the literature and identified gaps in our current knowledge of TSE, the following two chapters present studies that address these issues. Study 1 (Chapter 3) employs a meta-analysis to summarize what is known to date about the relation between teacher efficacy and indicators of effective classroom teaching. While early work on the relation between teacher efficacy and teaching behaviors gave us some indicators of specific behaviors of high efficacy teachers, the lack of a theoretical framework for effective classroom teaching resulted in lists of behaviors with no way to categorize and interpret findings. In that study, I distinguish between older conceptualizations of personal teaching efficacy and newer conceptualization of teacher self-efficacy. In addition, I assess

the extent to which TSE measures capture research-based strategies known to be effective for student learning and test this variable as a mediator of the relationship.

Study 2 (Chapter 4) is designed to provide depth to findings from Study 1 by uncovering how teachers' form their task-specific self-efficacy judgments in relation to their practices, student performance and their context. In one sense, this is a cognitive validity study to explore how consistent and aligned with research-based practices teacher interpretations of self-efficacy items are across domains of classroom management, student engagement and instructional strategies. This qualitative investigation uses think aloud methodology to capture teacher cognition as they respond to the most commonly used measure of TSE, the Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Hoy, 2001). Chapter 5 synthesizes and discusses findings from these two studies and makes recommendations for future research and conceptualization of TSE.

References

Abelson, A. G. (1997). Managing students with behavior disorders: Perceived efficacy of interventions. *Psychological Reports*, *80*(3, Pt 2), 1167–1170. http://doi.org/10.2466/pr0.1997.80.3c.1167

Allinder, R. M. (1994). The Relationship between Efficacy and the Instructional Practices of Special Education Teachers and Consultants. *Teacher Education and Special Education*, *17*(2), 86–95.

Allinder, R. M. (1995). An examination of the relationship between teacher efficacy and curriculum-based measurement and student achievement. *Remedial and Special Education*, *16*(4), 247–254. http://doi.org/10.1177/074193259501600408

Andersen, A. M., Dragsted, S., Evans, R. H., & S\orensen, H. (2004). The relationship between changes in teachers' self-efficacy beliefs and the science teaching environment of Danish first-year elementary teachers. *Journal of Science Teacher Education*, *15*(1), 25–38.

Armor, D., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., … Zellman, G. (1976). *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools.* Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED130243

Ashton, P. T. (1983). *A Study of Teachers' Sense of Efficacy. Final Report, Executive Summary.* (p. 43). Retrieved from http://search.proquest.com.proxy.lib.umich.edu/eric/docview/63515733/13E1E65B7463B5B9E44/34?accountid=14667

Ashton, P. T. (1984a). Teacher Efficacy: A Motivational Paradigm for Effective Teacher Education. *Journal of Teacher Education*, *35*(5), 28–32.

Ashton, P. T. (1984b). Teachers' Sense of Efficacy: A Self- or Norm-Referenced Construct? *Florida Journal of Educational Research*, *26*(1), 29–41.

Ashton, P. T., Olejnik, S., Crocker, L., & McAuliffe, M. (1982). Measurement Problems in the study of Teachers' Sense of Efficacy. Presented at the Annual Meeting of the American Educational Research Association.

Ashton, P. T., & Webb, R. B. (1986). *Making a difference: teachers' sense of efficacy and student achievement*. New York: Longman. Retrieved from http://hdl.handle.net/2027/[u]: mdp.39015010602426

Ball, D. L. (2000). Bridging practices intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, *51*(3), 241–247.

Ball, D. L., & Forzani, F. M. (2010). Teaching skillful teaching. *Educational Leadership*, *68*(4), 40–45.

Bandura, A. (n.d.). Teacher Self-Efficacy Scale. Unpublished manuscript.

Bandura, A. (1977a). Self efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, *84*(2), 191–215.

Bandura, A. (1977b). *Social learning theory*. Oxford England: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY US: W H Freeman/Times Books/ Henry Holt & Co.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development*, *67*(3), 1206–1222. http://doi.org/10.2307/1131888

Bautista, N. U. (2011). Investigating the Use of Vicarious and Mastery Experiences in Influencing Early Childhood Education Majors' Self-Efficacy Beliefs. *Journal of Science Teacher Education*, *22*(4), 333–349.

Behnke, P. (2007). *Brief in-service teacher training in a proactive approach to classroom behaviour management: Teacher and student outcomes*. ProQuest Information & Learning, US.

Berliner, D. C. (2004). Describing the Behavior and Documenting the Accomplishments of Expert Teachers. *Bulletin of Science, Technology & Society*, *24*(3), 200–212.

Berman, P., McLaughlin, M., Bass, G., Pauly, E., & Zellman, G. (1977). *Federal Programs Supporting Educational Change, Vol. VII: Factors Affecting Implementation and Continuation.* The Rand Corporation, 1700 Main Street, Santa Monica, California 90406 ($7.00). Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED140432

Blazevski, J. L. (2006). *Teacher efficacy for supporting student motivation*. ProQuest Information & Learning, US.

Bong, M. (2006). Asking the right question: How confident are you that you could successfully perform these tasks? In *Self-Efficacy Beliefs of Adolescents* (pp. 287–305). Information Age Publishing.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40. http://doi.org/10.1023/A:1021302408382

Borko, H., & Livingston, C. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Educational Research Journal*, *26*(4), 473–498. http://doi.org/10.2307/1162861

Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, *40*(2), 119–156.

Brophy, J. E. (2006). Observational Research on Generic Aspects of Classroom Teaching. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 755–780). Mahwah, NJ: Erlbaum.

Brophy, J. E., & Good, T. L. (1986). Teacher behaviors and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). New York: Macmillan.

Brouwers, A., & Tomic, W. (2000). Disruptive Student Behavior, Perceived Self-Efficacy, and Teacher Burnout. Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED450120

Brouwers, A., & Tomic, W. (2001). The factorial validity of scores on the Teacher Interpersonal Self-Efficacy Scale. *Educational and Psychological Measurement*, *61*(3), 433–445. http://doi.org/10.1177/00131640121971301

Brouwers, A., Tomic, W., & Stijnen, S. (2002). A confirmatory factor analysis of scores on the teacher efficacy scale. *Swiss Journal of Psychology*, *61*(4), 211–219. http://doi.org/10.1024//1421-0185.61.4.211

Bruce, C. D., Esmonde, I., Ross, J. A., Dookie, L., & Beatty, R. (2010). The effects of sustained classroom-embedded teacher professional learning on teacher efficacy and related student achievement. *Teaching and Teacher Education*, *26*(8), 1598–1608. http://doi.org/10.1016/j.tate.2010.06.011

Bursal, M. (2008). The impact of science methods courses on preservice elementary teachers' science teaching self-efficacy beliefs: Case studies from Turkey and the United States. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *68*(7-A).

Chong, W. H., Klassen, R. M., Huan, V. S., Wong, I., & Kates, A. D. (2010). The Relationships among School Types, Teacher Efficacy Beliefs, and Academic Climate: Perspective from Asian Middle Schools. *Journal of Educational Research*, *103*(3), 183–190.

Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education*, *28*(3), 258–277. http://doi.org/10.2307/749781

Cochran-Smith, M., McQuillan, P., Mitchell, K., Terrell, D. G., Barnatt, J., D'Souza, L., … Gleeson, A. M. (2012). A Longitudinal Study of Teaching Practice and Early Career Decisions: A Cautionary Tale. *American Educational Research Journal*, *49*(5), 844–880.

Coffey, J. E., Hammer, D., & Levin, D. M. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education*, *60*(2), 142+.

Cohen, D. K. (2010). Learning to Teach Nothing in Particular. *American Educator*, 44–54.

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, Instruction, and Research. *Educational Evaluation and Policy Analysis*, *25*(2), 119–142. http://doi.org/http://dx.doi.org.proxy.lib.umich.edu/10.3102/01623737025002119

Cohen, S. A., Hyman, J. S., Ashcroft, L., & Loveless, D. (1989). Comparing effects of metacognition, learning styles, and human attributes with alignment. In *Annual meeting of the American Educational Research Association*. San Francisco, CA.

Coladarci, T., & Breton, W. A. (1997). Teacher efficacy, supervision, and the special education resource-room teacher. *Journal of Educational Research*, *90*(4), 230–239.

Collmann, J. (2012). *Teacher stress and teacher-child interactions: An investigation of demographic, professional, and self-efficacy characteristics on the relationship between teacher stress and teacher-child interactions* (Unpublished doctoral dissertation). University of Virginia, US.

Corno, L. (2008). On Teaching Adaptively. *Educational Psychologist*, *43*(3), 161–173. http://doi.org/10.1080/00461520802178466

Cousins, J. B., & Walker, C. A. (2000). Predictors of educators' valuing of systematic inquiry in schools. *Canadian Journal of Program Evaluation*, *SPEC ISSUE*, 25–53.

Cripe, M. K. L. (2009). *A study of teachers' self-efficacy and outcome expectancy for science teaching throughout a science inquiry-based professional development program* (Unpublished doctoral dissertation). University of Akron, Ohio.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L., & Bransford, J. (2005). *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass, An Imprint of Wiley.

de Laat, J., & Watters, J. J. (1995). Science Teaching Self-Efficacy in a Primary School: A Case Study., 453–464.

Dolan, A. L., & McCaslin, M. (2008). Student perceptions of teacher support. *Teachers College Record*, *110*(11), 2423–2437.

Doyle, W. (1985). Recent Research on Classroom Management Implications for Teacher Education. *Journal of Teacher Education*, *36*(3), 31–35. http://doi.org/10.1177/002248718503600307

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*(4), 237–251. http://doi.org/10.3102/0013189X15584327

Duffin, L. C. (2010). *What do pre-service teachers know and believe about motivating students? a look at their motivational styles, strategies, and efficacy beliefs*. ProQuest Information & Learning, US.

Eccles, J. S., & Roeser, R. W. (1998). School and Community Influences on Human Development. In M. H. Borenstein & M. E. Lamb (Eds.), *Developmental Psychology: An advanced textbook* (4th ed., pp. 503–554). Mahwah, NJ: Erlbaum.

Edwards, J. L., & Green, K. E. (1999). *Persisters versus Nonpersisters: Characteristics of Teachers Who Stay in a Professional Development Program.* (p. 31). Retrieved from

http://search.proquest.com.proxy.lib.umich.edu/eric/docview/62389141/13E1FA187AD44263936/965?accountid=14667

Emmer, E. T., & Hickman, J. (1991). Teacher Efficacy in Classroom Management and Discipline. *Educational and Psychological Measurement*, *51*(3), 755 –765. http://doi.org/10.1177/0013164491513027

Enochs, L. G. (1993). *The Development and Partial Validation of Microcomputer Utilization in Teaching Efficacy Beliefs Instrument in a Science Setting.* (pp. 257–263). Retrieved from http://search.proquest.com.proxy.lib.umich.edu/eric/docview/62873554/13E1FA187AD44263936/1098?accountid=14667

Enochs, L. G., & Riggs, I. M. (1990). *Further Development of an Elementary Science Teaching Efficacy Belief Instrument: A Preservice Elementary Scale.* (p. 30). Retrieved from http://search.proquest.com.proxy.lib.umich.edu/eric/docview/63070103/13E1FA187AD44263936/1125?accountid=14667

Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the Mathematics Teaching Efficacy Beliefs Instrument. *School Science and Mathematics*, *100*(4), 194–202. http://doi.org/10.1111/j.1949-8594.2000.tb17256.x

Evertson, C. M., Anderson, C. W., Anderson, L. M., & Brophy, J. E. (1980). Relationships Between Classroom Behaviors and Student Outcomes in Junior High Mathematics and English Classes. *American Educational Research Journal*, *17*(1), 43–60. http://doi.org/10.3102/00028312017001043

Fahlman, M. M., Singleton, S. P., & Kliber, A. (2002). The Effects of Health Education Classes on Teaching Self-Efficacy in Preservice Teachers. *American Journal of Health Education*, *33*(2), 102–108.

Feldon, D. F. (2007a). Cognitive Load and Classroom Teaching: The Double-Edged Sword of Automaticity. *Educational Psychologist*, *42*(3), 123–137. http://doi.org/10.1080/00461520701416173

Feldon, D. F. (2007b). The implications of research on expertise for curriculum and pedagogy. *Educational Psychology Review*, *19*(2), 91–110. http://doi.org/10.1007/s10648-006-9009-0

Fives, H., & Buehl, M. M. (2009). Examining the factor structure of the Teachers' Sense of Efficacy Scale. *Journal of Experimental Education*, *78*(1), 118–134. http://doi.org/10.1080/00220970903224461

Gabriele, A. J., & Joram, E. (2007). Teachers' reflections on their reform-based teaching in mathematics: Implications for the development of teacher self-efficacy. *Action in Teacher Education*, *29*(3), 60–74.

Garcia, N., & Shaughnessy, M. (2015, October). *High Leverage Teaching Practices: What, Why, and How?* Presented at the Ohio Confederation of Teacher Education Organizations.

Gibson, S. (1983). *Teacher efficacy: A construct validation study* (Ph.D.). University of Southern California, United States -- California. Retrieved from http://search.proquest.com.proxy.lib.umich.edu/dissertations/docview/1629451605/abstract/872AFE0A573C4660PQ/2

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, *76*(4), 569–582. http://doi.org/10.1037/0022-0663.76.4.569

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining Teaching, Re-Imagining Teacher Education. *Teachers and Teaching: Theory and Practice*, *15*(2), 273–289.

Guo, Y., Connor, C. M., Yang, Y., Roehrig, A. D., & Morrison, F. J. (2012). The Effects of Teacher Qualification, Teacher Self-Efficacy, and Classroom Practices on Fifth Graders' Literacy Outcomes. *Elementary School Journal*, *113*(1), 3–24.

Guo, Y., Piasta, S. B., Justice, L. M., & Kaderavek, J. N. (2010). Relations among preschool teachers' self-efficacy, classroom quality, and children's language and literacy gains. *Teaching and Teacher Education*, *26*(4), 1094–1103. http://doi.org/10.1016/j.tate.2009.11.005

Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, *31*(3), 627–643. http://doi.org/10.2307/1163230

Hardré, P. L., & Sullivan, D. W. (2008). Teacher perceptions and individual differences: How they influence rural teachers' motivating strategies. *Teaching and Teacher Education*, *24*(8), 2059–2075. http://doi.org/10.1016/j.tate.2008.04.007

Hardré, P. L., & Sullivan, D. W. (2009). Motivating Adolescents: High School Teachers' Perceptions and Classroom Practices. *Teacher Development*, *13*(1), 1–16.

Heneman, H. G., Kimball, S., & Milanowski, A. (2006). *The Teacher Sense of Efficacy Scale: Validation Evidence and Behavioral Prediction. WCER Working Paper No.*

2006-7 (p. 21). Wisconsin Center for Education Research. School of Education; University of Wisconsin-Madison, 1025 West Johnson Street Suite 785, Madison, WI 53706. Retrieved from http://search.proquest.com.proxy.lib.umich.edu/eric/docview/62036497/13E1E65B7463B5B9E44/31?accountid=14667

Henson, R. K. (2001). Teacher Self-Efficacy: Substantive Implications and Measurement Dilemmas. Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED452208

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107.

Holzberger, D., Philipp, A., & Kunter, M. (2013). How Teachers' Self-Efficacy Is Related to Instructional Quality: A Longitudinal Analysis. *Journal of Educational Psychology*. http://doi.org/10.1037/a0032198

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal*, *38*(3), 499.

Jackson, P. W. (1990). *Life in Classrooms*. Teachers College Press.

Jamil, F. M., Downer, J. T., & Pianta, R. C. (2012). Association of Pre-Service Teachers' Performance, Personality, and Beliefs with Teacher Self-Efficacy at Program Completion. *Teacher Education Quarterly*, *39*(4), 119–138.

Kazemi, E. (1998). Discourse That Promotes Conceptual Understanding. *Teaching Children Mathematics*, *4*(7), 410–14.

Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y. F., & Georgiou, T. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemporary Educational Psychology*, *34*(1), 67–76. http://doi.org/10.1016/j.cedpsych.2008.08.001

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, *102*(3), 741.

Klassen, R. M., Durksen, T. L., & Tze, V. M. C. (2014). Teachers' self-efficacy beliefs: Ready to move from theory to practice? In P. W. Richardson, S. A. Karabenick, & H. M. G. Watt (Eds.), *Teacher motivation: Implications for theory and practice* (pp. 100–115). U.S.: Routledge/Taylor & Francis.

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher Efficacy Research 1998–2009: Signs of Progress or Unfulfilled Promise? *Educational Psychology Review*, *23*(1), 21–43. http://doi.org/10.1007/s10648-010-9141-8

Klieme, E. (2012, August). *Qualities and effects of teaching: Towards a conceptual theory of teaching.* Presented at the EARLI SIG meeting, Educational Effectiveness, Zurich, Switzerland.

Knoblauch, D., & Chase, M. A. (2015). Rural, suburban, and urban schools: The impact of school setting on the efficacy beliefs and attributions of student teachers. *Teaching and Teacher Education*, *45*, 104–114. http://doi.org/10.1016/j.tate.2014.10.001

Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York: Holt, Rinehart and Winston. Retrieved from http://hdl.handle.net/2027/[u]: mdp.39015004798677

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. http://doi.org/10.1207/s15430421tip4104_2

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, *9*(3), 231–251. http://doi.org/10.1007/s10984-006-9015-7

Labone, E. (2004). Teacher Efficacy: Maturing the Construct through Research in Alternative Paradigms. *Teaching & Teacher Education*, *20*(4), 341–359.

Lampert, M., & Clark, C. M. (1990). Expert knowledge and expert thinking in teaching: A reply to Floden and Klinzig. *Educational Researcher*, *19*(4), 21–23, 42.

Lauermann, F., & Karabenick, S. A. (2011). Taking teacher responsibility into account(ability): Explicating its multiple components and theoretical status. *Educational Psychologist*, *46*(2), 122–140. http://doi.org/10.1080/00461520.2011.558818

Lauermann, F., & Karabenick, S. A. (2013). The meaning and measure of teachers' sense of responsibility for educational outcomes. *Teaching and Teacher Education*, *30*, 13–26. http://doi.org/10.1016/j.tate.2012.10.001

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and Extrinsic Motivational Orientations in the Classroom: Age Differences and Academic Correlates. *Journal of Educational Psychology*, *97*(2), 184–196. http://doi.org/10.1037/0022-0663.97.2.184

Malow-Iroff, M. S., O'Connor, E. A., & Bisland, B. M. (2004). *Alternatively Certified Teachers: Efficacy Beliefs and Ideology* (p. 8). Retrieved from

http://search.proquest.com.proxy.lib.umich.edu/eric/docview/62075057/13E1F6C05C86FE37B34/805?accountid=14667

Mavropoulou, S., & Padeliadu, S. (2002). Teachers' causal attributions for behaviour problems in relation to perceptions of control. *Educational Psychology*, *22*(2), 191–202. http://doi.org/10.1080/01443410120115256

Mulholland, J., & Wallace, J. (2001). Teacher Induction and Elementary Science Teaching: Enhancing Self-Efficacy. *Teaching and Teacher Education*, *17*(2), 243–61.

Newmann, F. M. (1996). *Authentic achievement: restructuring schools for intellectual quality*. San Francisco: Jossey-Bass Publishers. Retrieved from http://hdl.handle.net/2027/[u]: mdp.39015038157007

Nie, Y., Tan, G. H., Liau, A. K., Lau, S., & Chua, B. L. (2013). The Roles of Teacher Efficacy in Instructional Innovation: Its Predictive Relations to Constructivist and Didactic Instruction. *Educational Research for Policy and Practice*, *12*(1), 67–77.

O'Neill, S. C., & Stephenson, J. (2011). The measurement of classroom management self-efficacy: A review of measurement instrument development and influences. *Educational Psychology*, *31*(3), 261–299. http://doi.org/10.1080/01443410.2010.545344

Pfitzner-Eden, F., Thiel, F., & Horsley, J. (2014). An adapted measure of teacher self-efficacy for preservice teachers: Exploring its validity across two countries. *Zeitschrift Für Pädagogische Psychologie / German Journal of Educational Psychology*, *28*(3), 83–92. http://doi.org/10.1024/1010-0652/a000125

Pianta, R. C., Belsky, J., Vandergrift, N., Houts, R., & Morrison, F. J. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal*, *45*(2), 365–397. http://doi.org/10.3102/0002831207308230

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity. *Educational Researcher*, *38*(2), 109–119. http://doi.org/10.3102/0013189X09332374

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3*. Baltimore, MD US: Paul H Brookes Publishing.

Pianta, R. C., La Paro, K. M., Payne, C., Cox, M. J., & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school characteristics and child outcomes. *The Elementary School Journal*, *102*(3), 225–238. http://doi.org/10.1086/499701

Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-Level Positive Behavior Supports in Schools Implementing SW-PBIS: Identifying Areas for Enhancement. *Journal of Positive Behavior Interventions*, *15*(1), 39–50. http://doi.org/10.1177/1098300712459079

Rohrkemper, M., & Corno, L. (1988). Success and failure on classroom tasks: Adaptive learning and classroom teaching. *The Elementary School Journal*, *88*(3), 297–312. http://doi.org/10.1086/461540

Rohs, J. M. (2007). *Child outcomes in head start classrooms: The impact of teacher beliefs and interactions*. University of Missouri-Kansas City, US.

Ryan, A. M., Kuusinen, C. M., & Bedoya-Skoog, A. (2015). Managing peer relations: A dimension of teacher self-efficacy that varies between elementary and middle school teachers and is associated with observed classroom quality. *Contemporary Educational Psychology*, *41*, 147–156. http://doi.org/10.1016/j.cedpsych.2015.01.002

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. http://doi.org/10.1006/ceps.1999.1020

Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences Among Teachers in a Task Characterized by Simultaneity, Multidimensional, and Immediacy. *American Educational Research Journal*, *28*(1), 63–88. http://doi.org/http://dx.doi.org.proxy.lib.umich.edu/10.3102/00028312028001063

Schmitz, G. S., & Schwarzer, R. (2002). Individuelle und kollektive Selbstwirksamkeitserwartung von Lehren / Teachers' individual and collective self-efficacy. *Zeitschrift Für Pädagogik*, *44*, 192–214.

Schunk, D., & Pajares, F. (2009). Self-Efficacy Theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 35–53). New York: Routledge.

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*(3), 611–625. http://doi.org/10.1037/0022-0663.99.3.611

Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, *26*(4), 1059–1069. http://doi.org/10.1016/j.tate.2009.11.001

Skinner, E. A. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology*, *71*(3), 549–570. http://doi.org/10.1037/0022-3514.71.3.549

Solomon, D., Battistich, V., Watson, M., Schaps, E., & Lewis, C. (2000). A six-district study of educational change: Direct and mediated effects of the child development project. *Social Psychology of Education*, *4*(1), 3–51. http://doi.org/10.1023/A:1009609606692

Soodak, L. C., & Podell, D. M. (1996). Teacher efficacy: Toward the understanding of a multi-faceted construct. *Teaching and Teacher Education*, *12*(4), 401–411. http://doi.org/10.1016/0742-051X(95)00047-N

Thompson, J., Windschitl, M., & Braaten, M. (2013). Developing a Theory of Ambitious Early-Career Teacher Practice. *American Educational Research Journal*, *50*(3), 574–615. http://doi.org/10.3102/0002831213476334

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. http://doi.org/10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*(2), 202–248. http://doi.org/10.2307/1170754

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*(4), 548–573. http://doi.org/10.1037/0033-295X.92.4.548

Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, *73*(1), 287–301. http://doi.org/10.1111/1467-8624.00406

Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education*, *21*(7), 747–766. http://doi.org/10.1016/j.tate.2005.05.009

Williams, W. M., Blythe, T., White, N., Li, J., Gardner, H., & Sternberg, R. J. (2002). Practical intelligence for school: Developing metacognitive sources of achievement in adolescence. *Developmental Review*, *22*(2), 162–210. http://doi.org/10.1006/drev.2002.0544

Woolfolk, A. E., & Hoy, W. K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology*, *82*(1), 81–91. http://doi.org/10.1037/0022-0663.82.1.81

Wu, S. C., & Chang, Y. L. (2006). Elementary Teacher Education and Teacher Efficacy Toward Mathematics and Science. In J. Novotna, H. Moraova, M. Kratka, & N. Stehlikova (Eds.), *Pme 30: Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education, Vol 5* (pp. 417–424). Praha 1: Int Grp Psychol Math Educ.

Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education*, *45*, 128–136. http://doi.org/10.1016/j.tate.2014.10.002

CHAPTER 3

A META-ANALYTIC REVIEW OF TEACHER SELF-EFFICACY AND EFFECTIVE CLASSROOM TEACHING PRACTICES

Introduction

Teacher self-efficacy (TSE) is important for teachers' psychological well-being, including their feelings about their job satisfaction, intent to stay in the profession, and professional commitment (Coladarci, 1992; Skaalvik & Skaalvik, 2007a). Moreover, teachers with higher levels of TSE have more positive attitudes about and higher goals for students (Gibson & Dembo, 1984), persist more with those who are struggling, and are less controlling (Enochs, Scharmann, & Riggs, 1995). The relationship between TSE and teachers' use of research-based strategies known to be effective for student learning, however, is less clear. Early studies focused on the types of practices associated with high self-efficacy teachers, such as whole group versus small-group instruction (Gibson & Dembo, 1984); yet the appropriateness of these strategies or the quality of such practices was not a focus of the investigations. A recent meta-analysis partially addressed this gap in the literature by showing that teacher self-efficacy was related to overall teaching quality (Klassen & Tze, 2014); however, no clear criteria were used to define teaching quality across studies. *Clearly identifying for which types of teaching behavior teacher self-efficacy matters would provide needed insight to move this construct from theory to practical application* (Klassen, Durksen, & Tze, 2014).

Accordingly, a more complete meta-analysis is required that capitalizes on recent advances in the conceptualization and assessment of effective classroom teaching. Specifically, the present meta-analysis investigated the extent to which there is evidence for: (a) the overall relationship between teacher self-efficacy and effective (research-based) teaching practices in academic classroom settings, (b) the strength of the association between teacher self-efficacy and three core domains of effective classroom teaching (structured classroom management, supportive classroom climate, and cognitive activation) (Baumert et al., 2010; Klieme, 2012), and (c) the potential moderating effect of different operationalizations of teacher self-efficacy on the association between teacher self-efficacy and teaching quality.

Literature Review

Multi-dimensionality of teacher self-efficacy and teaching behavior

As reviewed in Chapter 2, TSE was originally conceptualized and measured in line with locus of control theory (Armor et al., 1976) in which teachers evaluated their belief they could have a positive impact on student motivation and performance above that of external, negative factors (Ashton, 1984a; Ashton & Webb, 1986). Current conceptualizations of TSE, used in the present study, are based in social cognitive theory that assesses future-oriented beliefs about whether teachers can execute courses of action at a desired level of competence (Bandura, 1997). In addition, teacher self-efficacy beliefs are self-assessments of capabilities to perform specific tasks rather than global assessments of functioning (Bong & Skaalvik, 2003). Accordingly, teachers' sense of efficacy is typically assessed in three task domains that are common across international measures of teacher

self-efficacy: classroom management, student engagement, and instructional strategies (Skaalvik & Skaalvik, 2007a; Tschannen-Moran & Hoy, 2001).

Classroom management efficacy represents teachers' self-efficacy for responding to disruptive behavior, monitoring work, and establishing norms and routines (Emmer & Hickman, 1991). *Student engagement efficacy* captures teachers' beliefs they can effectively motivate students, engage parents, and help students value learning (Blazevski, 2006). *Instructional strategies efficacy* refers to teachers' beliefs that they can perform such instructional tasks as adapting content to student needs, questioning students, and assessing student learning (Tschannen-Moran & Hoy, 2001). While other domains of teacher self-efficacy have been identified (e.g., managing peer relations; Ryan, Kuusinen, & Bedoya-Skoog, 2015), the three domains in the present analysis focuses most closely on teachers' daily work in the classroom with students around academic content. Specifically, the goal is to identify the types of effective teaching behaviors most often demonstrated by high efficacy teachers.

The following section describes three domains of effective classroom teaching represented by current research: a) structured classrooms and instructional management, b) supportive classroom climate, and c) cognitive activation (Brophy, 2006; Evertson, Anderson, Anderson, & Brophy, 1980; Holzberger, Philipp, & Kunter, 2013; Klieme, 2012; Pianta & Hamre, 2009). These three teaching quality domains generally correspond to the three domains of teacher self-efficacy research described above (Klieme, 2012). Specifically, I investigate the extent that (a) teachers' sense of self-efficacy in any of the three domains is related broadly to observations of teaching practices in any domain, (b) TSE, in any domain, is related to each separate domain of effective teaching practice, and

(c) publication status, type of self-efficacy measures, and item construction moderate these relations.

Three domains of effective classroom teaching

Effective classroom teaching is defined as the use of strategies that have been associated with student learning gains in core academic content areas (specifically, mathematics, English, science and history). For each of the three domains of effective classroom teaching, I first describe its conceptualization and compare it to the corresponding conceptualization of teacher self-efficacy for that domain in order to assess the extent of specificity and domain matching (McWilliams, Nier, & Singer, 2013). The specificity matching principle states that specific predictors best predict domain-specific behaviors rather than broad behaviors. A lack of coherence between TSE and classroom teaching may limit the ability to detect associations between the two constructs.

S*tructured classroom management* describes a classroom environment that runs smoothly and efficiently because teachers move students towards specific learning goals set by teachers and communicated effectively to students. This domain of effective teaching was one of the first to be studied in connection to student achievement (Evertson et al., 1980). Teachers who score highly in this domain were observed to maintain a brisk pace in their classroom that is rarely derailed by students' behavioral issues. They proactively manage the classroom by establishing routines and procedures, cue students as to situational expectations, and consistently enforce rules (Brophy & Good, 1986). Supporting this ability to structure and manage the classroom, teachers present content in a coherent and logical manner, organize resources efficiently, and assessments are aligned with learning goals (Cohen, Hyman, Ashcroft, & Loveless, 1989).

The comparable assessment of TSE, however, does not fully capture the later aspect of this domain. Rather, recent reviews of classroom management efficacy have found that classroom management scales tend to emphasize behavior management only and may not capture current psychological and educational research on effective classroom management (O'Neill & Stephenson, 2011). Other domains of TSE, such as instruction, may capture elements of teachers' perceived abilities to make their expectations clear, though this is not consistent across all scales of TSE (e.g., cf. Skaalvik & Skaalvik, 2010; Tschannen-Moran & Hoy, 2001; Tschannen-Moran & Johnson, 2011). Therefore, the strength of the associations between TSE and structured classroom management may be weak to moderate, depending on the specific operationalization of self-efficacy and the degree to which classroom management is limited to the dimension of behavioral management versus includes other facets as well.

Supportive classroom climate represents the extent to which teaching is in tune with student needs for a safe environment for learning. The relationships between teachers and students in the classroom has long been studied, and beyond achievement has also been positively associated with student self-concept, self-efficacy, engagement and student expectations for success (Dolan & McCaslin, 2008; Eccles & Roeser, 1998; Wentzel, 2002). Early studies focused mainly on teacher actions that conveyed a sense of warmth, such as frequent smiling (Kounin, 1970), enthusiasm, and use of praise rather than criticism (Evertson et al., 1980). This research shows the importance of teachers' expectations for children (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996), and to what degree teachers view individual differences as opportunities for learning rather than obstacles to overcome (Corno, 2008). Also important is the degree to which teachers are

autonomy supportive, or able to adapt instruction to students' personal needs and individual differences (Pianta, La Paro, Payne, Cox, & Bradley, 2002). Such support can help to create a safe climate for learning in which students interact with each other respectfully, cooperate, and see each other as resources for learning (Pianta & Hamre, 2009; Rohrkemper & Corno, 1988).

Teacher self-efficacy captures most, but not all, of these aspects of a supportive climate. TSE for student engagement typically assesses teachers' perceived capabilities to influence individual students' intrinsic motivation to do well in schoolwork or value learning. While TSE for student engagement may not fully assess teachers' capability to create a supportive environment among and between students, elements of TSE for classroom management and instruction may assess the ability to work together cooperatively or differentiate instruction based on student needs (e.g., the long version of TSES). The degree of association between self-efficacy and teaching quality would depend on the degree to which scales assess these various aspects of a supportive climate. Therefore, expected associations between TSE and supportive climate domain may be weak to moderate because of some lack of congruity between constructs. Moreover, teachers may feel efficacious to motivate individual students yet less able of create an environment in which students work together cooperatively.

Cognitive activation represents teachers' provision of supports for cognitive engagement with appropriately challenging learning activities, particularly those that require students to construct meaning through discourse with the teacher and other students (Boston & Smith, 2009; Holzberger et al., 2013). Educational psychologists typically use the phrase *cognitive demand* to refer to the types of thinking students are asked to do that

can be assessed through teacher questions, activities and assessments (Boston & Smith, 2009; Holzberger et al., 2013). The role of thoughtful discourse around content has been conceptualized as the opportunities teachers provide for students to answer questions that demand higher level thinking (Cobb, Boufi, McClain, & Whitenack, 1997; Kazemi, 1998). Instruction in metacognitive strategies has been shown to strongly relate to student achievement (Williams et al., 2002), particularly the extent to which teachers emphasize the processes used to generate answers rather than just correct answers (Darling-Hammond & Bransford, 2005). This has often been investigated through the use of authentic problems that require students to solve challenging problems (Kazemi, 1998; Newmann, 1996).

This domain of instruction is less often seen in the assessment of teacher self-efficacy. TSE scales may capture teachers' perceived abilities to challenge students and get students to think critically (e.g., long version of TSES), yet the role of thoughtful discourse and emphasizing process over answer generation is rarely assessed, which is related to cognitive demand but yet it may not capture use of metacognitive strategies and focus on the process of learning. Thus, the associations between TSE and cognitive activation may be weak, particularly in comparison to other domains.

In sum, current approaches to the assessment of TSE capture some but not all of the practices associated with contemporary and research-based views about what effective teachers do in the classroom (O'Neill & Stephenson, 2011). The following section reviews the extant literature on the relationship between teacher self-efficacy and teaching quality and describes how the present study extends that work.

Prior reviews of teacher self-efficacy and teaching quality

Early qualitative research identified specific, isolated strategies that high efficacy teachers were observed to use in the classroom. More efficacious teachers in special education were found to set higher goals for students and were willing to persist with struggling students (Allinder, 1995; Ashton & Webb, 1986). This work, however, was grounded in older conceptualizations of teacher self-efficacy described previously. Moreover, numerous scholars have identified conceptual and measurement problems with early research on TSE (Brouwers, Tomic, & Stijnen, 2002; Guskey & Passaro, 1994; Woolfolk & Hoy, 1990). Yet, current empirical and conceptual reviews have not accounted for the different measures and conceptualization of teacher self-efficacy as potential moderators of examined relationships (Klassen & Tze, 2014; O'Neill & Stephenson, 2011). Accordingly, one goal of the present study is to assess the degree to which the strength of the relationship between teaching practice and teacher self-efficacy is moderated by the operationalization of teacher self-efficacy.

The recent meta-analysis of teacher self-efficacy (Klassen & Tze, 2014) and its relationship with teaching quality identified eight studies in which teaching quality had been assessed by an external observer such as a supervisor, researcher or student. Using six of these studies, the authors found a moderate positive correlation with teacher self-efficacy and ratings of teacher performance ($r = .28$). This association demonstrates that there is a moderate but significant relationship between teachers' beliefs they can accomplish teaching tasks and the quality of their instruction as rated by external observers. However, certain aspects of that meta-analysis must be considered before drawing conclusions about the robustness of this finding. First, studies of teachers in such non-

academic subjects as music or physical education were included (e.g., Barnes, 1998). Teaching in these domains is qualitatively different than teaching in academic content areas because greater emphasis is placed on performance-based instruction of music or physical tasks rather than the instruction in concepts, ideas and critical thinking in academic domains. Second, studies of teachers at university levels (Balam, 2006) were included with pre-K-12 teachers, although the conditions for teaching and expectations for performance in these settings can differ substantially. Third, studies in which TSE for tasks other than classroom teaching (e.g., TSE for computer use; Holden & Rada, 2011) were included which, according to the specificity-matching principle, may influence the true association between the predictor and outcome. In addition, no distinction was made between studies that used different conceptualizations and corresponding measures of TSE. Finally, lack of clear criteria for quality of instruction limit the ability to understand for which aspects of instruction TSE is most important or if it is equally important for all domains of instruction. Therefore, the present study focused exclusively on studies of self-efficacy for teaching in classrooms among preK-12 teachers of academic subjects.

The importance of item construction

Constructing valid and reliable measures of self-efficacy is a challenge that requires one to decide the right level of specificity at which to measure the task (Bandura, 2006; Bong, 2006). Past measures of teacher self-efficacy were critiqued for using a past-oriented or attribution-focused assessment of teaching effectiveness (see Chapter 2 for a comprehensive review). These critiques were taken to heart when constructing the Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Hoy, 2001), in which all items are formed as questions that asks teachers "how much they can do" or "how well

they can" accomplish specific tasks. Accordingly, the present study accounted for whether TSE was assessed using a future-oriented assessment of capabilities.

Also important and less addressed is the consideration of obstacles or gradations of challenges in self-efficacy items (Bandura, 2006). These obstacles allow teachers to make appraisals that reflect the level of difficulty they believe they can surmount. Without challenges in the items, teachers may think of varying levels of difficulty of the item, casting doubt on the comparability of teachers' self-efficacy ratings. Moreover, some evidence exists that teachers feel less efficacious to solve problems, particularly those related to managing disruptive behavior (Chang, 2013). Therefore, the present meta-analysis also categorized each TSE scale by the percentage of items that included an obstacle, challenge or problem to be surmounted.

Finally, specificity matching, as discussed above, is also a concern. Self-efficacy scales "must be tailored to the particular domain of functioning that is the object of interest" in order to maximize the validity of the measure (Bandura, 2006, p.308). As some scales are a mix of items assessing classroom management and instruction (c.f. Enochs et al., 1995; Gibson & Dembo, 1984), each TSE scale was categorized by percentage of items that assessed the domain of teaching being investigated.

Goals of the present meta-analysis

The present meta-analysis builds on prior knowledge in a number of ways. First, it focuses exclusively on elementary and secondary teachers of academic content areas: English, mathematics, science, and history. Second, studies that measured TSE for tasks other than direct classroom teaching were excluded in order to draw robust conclusions regarding the relation between self-efficacy for teaching and observed quality of teaching.

Third, characteristics of teacher self-efficacy important for validity (specifically, the extent to which items are future-oriented judgments of capability and whether a challenge or obstacle is present in the item) were used as potential moderators of the relationship to determine whether more conceptually sound social-cognitive measures (e.g., the TSES) or those scales that included more items with challenges or obstacles or items that matched the domain of teaching demonstrate stronger associations with teaching effectiveness. Finally, the present meta-analysis extends prior knowledge by determining the domains of effective classroom teaching that are most strongly related to teacher self-efficacy. Such a rigorous and comprehensive approach addresses recent calls to advance teacher self-efficacy research from theory to practice by linking teachers' self-efficacy beliefs with selected dimensions of teaching quality (Klassen et al., 2014).

Method

Sample and selection criteria. Three education and psychology databases (PsycInfo, Eric ProQuest, and Web of Science) were searched for combinations of the words "teacher(s)," "teaching" and "efficacy" within the title or abstract. This approach is more comprehensive than past analyses that have only searched within the titles of studies (e.g., Klassen & Tze, 2014). Prior meta-analyses on teacher self-efficacy were also searched to identify additional work that may have been omitted from our search. This procedure yielded 2932 potential studies. Abstracts were read by trained research assistants and myself for the following inclusion criteria: 1) use of a quantitative measure of individual teacher self-efficacy for teaching (e.g., not collective self-efficacy or self-efficacy for computer use); 2) participants teaching in preK-12 classrooms; 3) teachers of English, science, mathematics or social studies/history; 4) observation of teaching by

external observer (e.g., researcher, supervisor or student), and 5) study written in English. Inter-rater agreement was good: 85% agreement between coder 1 and myself and 81% with coder 2 and myself (coders did not read the same studies). All disagreements were resolved through discussion.

Because of the large number of studies included in this initial search, additional measures were taken to ensure that reliable identification of relevant studies. The abstracts and titles of the initial 2932 studies (stored in Zotero bibliographic management system) were searched again for terms indicating observation of teaching in the classroom: "observation," "observe," "teaching practice," "teacher quality," and "classroom quality." An additional two studies that met inclusion criteria were identified and coded (Jamil, Downer & Pianta, 2012; Pakarinen et al., 2010). Finally, the final pool was cross-referenced with the list of studies identified by Klassen & Tze (2014), excluding studies in which the sample was not academic teachers or did not address self-efficacy for teaching. This resulted in a sample of 40 studies, which is a substantially larger sample pool than the number of studies included in past meta-analyses of teacher self-efficacy and teaching quality (e.g., Klassen & Tze, 2014).

Study Coding. All of the studies were coded for zero-order Pearson correlation coefficients (r) that were either reported in published research, sent by the author(s), calculated from raw data sent by the author(s), or obtained from dissertations. Also coded were study characteristics (e.g., teacher characteristics, teacher self-efficacy measure used, sampling procedures), and which domain(s) of effective classroom teaching were represented by the effect size. Table 5 presents the criteria used to categorize domains of effective teaching based on extensive review of the literature in each domain.

For each domain, three subcategories were identified through synthesis of the available literature on effective classroom teaching (Brophy, 2006; Klieme, 2012; Pianta et al., 2002). *Structured classroom management* is comprised of 1) opportunity to learn, 2) proactive behavior management, and 3) instructional coherence. *Supportive classroom climate* was operationalized as the assessment of positive 1) teacher-student interactions, 2) student-student interactions, and 3) cooperative structure of activities. Cognitive activation consisted of 1) thoughtful discourse, 2) metacognitive strategies, and 3) problem-solving. These subcategories were generated in order to assist in making coding decisions. An effect size was considered to represent a domain of effective instruction (the DV) if at least 33% of the items in the observational tool could be categorized within that domain (any subcategory). I chose 33% as a cutoff in order to account for studies that used effect sizes from studies that used three-dimensional holistic observation measures that corresponded with the three domains of effective instruction.

Moderator Coding. Two different approaches were applied to coding teacher self-efficacy measures. First, scales were broadly categorized for whether they represented a social-cognitive conceptualization of teacher self-efficacy ("I can" or "am capable" represented by measures such as the Teachers' Sense of Efficacy Scale; Tschannen-Moran & Hoy, 2001) rather than past functioning or attributions for student outcomes (represented by measures such as the Teacher Efficacy Scale; Gibson & Dembo, 1984). This resulted in a binary moderator where 1=social-cognitive perspective and 0=other perspective. All studies were coded by two coders with 78% agreement; disagreements were resolved through discussion until there was 100% agreement.

Second, I coded individual items from each TSE measure on three dimensions: 1) whether item was constructed as a future-oriented assessment of capabilities using "can" "able" or "capable," 2) whether the item contained an obstacle, challenge or problem that the teacher needed to solve, and 3) whether and which of the three domains of effective teaching described previously was captured by each item. The number of items that received each code were summed and divided by the total number of items on the scale, then multiplied by 100 consistent with methods used in previous meta-analyses (Hulleman, Schrager, Bodmann, & Harackiewicz, 2010). This resulted in five continuous moderator variables, three of which were the percentage of items per scale that represented a particular domain of effective teaching (e.g., structured classroom management), and the remaining two were the percentage of items with a future-oriented assessment of capabilities, as well as the percentage of items including an explicit reference to a challenge, obstacle, or problem.

Finally, studies were coded as published research or not to assess for potential publication bias. Unpublished studies included dissertations (k=14), conference papers (k=2), and research reports (k=1).

Calculating effect sizes. Studies that provided more than one effect size were averaged to meet the assumption of independence (Lipsey & Wilson, 2001). Correlation coefficients were transformed to Fisher's z using the *metafor* package of R statistical software (Viechtbauer, 2010). Each effect size was weighed by the inverse variance in order to give more weight to studies with larger sample sizes. One study was subsequently excluded from the analyses due to its low sample size (n = 4) for which sampling variances could not be calculated, resulting in a final sample size of 39. Additionally, I performed an

exclusion sensitivity analysis to identify potential outliers (Lipsey & Wilson, 2001), systematically excluding one study at a time. Since the overall coefficient remained statistically significant in each model, all 39 studies were included in the final model.

Results

Descriptives. Tables 6 and 7 present the results of coding for type of measure used and observational tool. While only one study used a unique measure of TSE designed specifically for the particular study (e.g., no published validity study for this measure), half of the studies (k = 20) used a unique observational measure. Table 8 presents the extent to which the 40 studies (before exclusion of one study) showed domain matching between teacher self-efficacy and teaching effectiveness measures (e.g., classroom management efficacy and its relation to structured classroom management), shown on the diagonal. Few studies (30%) exactly matched domains of teacher self-efficacy with corresponding domains of effective teaching (k =12). More commonly, studies assessed more than one domain of teaching effectiveness along with multi-dimensional teacher self-efficacy scales based in social cognitive theory (k = 14) or older conceptualizations of general and personal teaching efficacy (k = 14).

[Insert Tables 5, 6 and 7 about here]

Figure 5 displays results of the moderator coding for the two most frequently used measures of TSE, the Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Hoy, 2001) and the Teacher Efficacy Scale (TES; Gibson & Dembo, 1984). The long and short version of the TSES contain 100% of items with a future-orientation, compared to 18% of items on the TES (shown by subscale of personal teaching efficacy and entire scale with

general teaching efficacy measure). The TES scale, however, contained a greater number of items with a problem or challenge (71%) mostly located within the PTE subscale (44%). The PTE subscale also had more items concerning cognitive activation (78%) compared to any other scale and the entire TES had more items concerning supportive climate (71%) compared to any other measure. Comparing the short, 12-item version and long, 24-item versions of the TSES, the long version contained more items coded as structured management (33%) compared to the short version (15%). Figure 6 displays the reason for this difference by showing coding by subscale of the TSES, separated by short or long version. Items categorized by the authors as "TSE for student engagement" on the long version of the TSES were coded as cognitive activation (e.g., "To what extent can you help your students to think critically?"), and some items categorized by the authors as "classroom management" were coded as supportive climate based on findings from the subsequent chapter (e.g., "To what extent can you calm a disruptive student?").

[Insert Figures 5 and 6 about here]

The 40 studies represent 137 effect sizes (133 effect sizes with the exclusion of one study), summarized in Table 9, which also includes the number of participants, measure of teacher self-efficacy used, observational measure used, and the domain(s) of effective teaching represented by the measure. Effect sizes are reported in terms of Pearson r correlations before z transformations.

[Insert Table 9 about here]

It should be noted that most of the participants in these studies were in-service teachers (k = 31) at the elementary level (k = 26). Six studies included teachers of mixed

grades and six of secondary teachers only. The sample sizes ranged from 8 to 896, and average years of teaching experience of participants ranged from 0 to 22, with four studies not reporting the average years of teaching experience.

Statistical modeling. A random effects model with restricted maximum-likelihood estimator was chosen, which accounts for heterogeneity of effects between studies and permits generalizing results beyond the set of studies included (Lipsey & Wilson, 2001). Results of Cochran's Q-test ($Q(df = 38) = 194.73$, $p < .001$) confirmed that the variability in effect sizes was larger than would be expected based on sampling variability alone. I then tested four statistical models: the full model tested the overall strength of the relationship between teacher self-efficacy and effective teaching practices in any domain. Subsequent models tested the association between teacher self-efficacy and each domain of effective teaching. Finally I tested for potential moderators of each of these models. All effect sizes were transformed back to Pearson's r correlation for ease of interpretation.

The overall effect size between teacher self-efficacy and effective teaching practices (in any domain) was small according to Cohen (1988) criteria where .2 indicates a small effect size, .5 indicates a medium effect and .8 indicates a large effect size. Although statistically significant, is smaller than that found in previous meta-analyses, r = .17, SE = 0.05, z = 3.51, 95% CI=[0.08, 0.27], p=.001. Figure 7 displays the forest plot of all effect sizes. Three subsequent models that tested the association between teacher self-efficacy and teaching practice in each of the three domains of teaching (structured management, climate and cognitive activation) are shown in Table 10. The effect size for the 21 studies included in the Structured Management model was positive, but not significant, even though the lower bound of the 95% CI approximated zero, r = .11, SE =

0.06, z = 1.85, 95% CI=[0.00, 0.22], *p*=.07. However, the effect sizes were significant for the 14 studies included in the Supportive Classroom Climate model, r = .17, SE = 0.05, z = 3.72, 95% CI=[0.08, 0.27], p= .0002, and the 18 studies included in the Cognitive Activation model, r = .17, SE = 0.06, z = 2.77, 95% CI=[0.05, 0.29], p < .006.

[Insert Figure 7 about here]

Moderator analyses. Moderator analyses were conducted to assess whether the identified heterogeneity in the recorded correlations across studies could be explained by the following potential moderators: measure of TSE (based on social-cognitive theory vs. not), publication status (published research vs. not), percentage of items per scale that were future-oriented assesment of capabilities, and percentage of items that included a challenge or a problem, and percentage of TSE that matched the domain of effective teaching in the model. An ANOVA-like mixed effects model was used for the first two analyses because these moderator variables are categorical, rather than continuous. For the continuous variables, a meta-regression was used instead. The overall Q statistic was partitioned into two types of variances: Q_M, or the omnibus test of between-group variance in effect size and Q_W, the within-groups variance in effect size. A non-significant between-groups variance indicates that the moderator does not account for significant variability in effect size. The Q_W statistic was not examined if the QM statistic was not significant. As discussed in the following sections, the hypothesized moderators did not have significant effects.

Analyses of Moderator Effects in the Effective Teaching Model

Average effect sizes for the TES scale were .21 (p=.004) for studies using a social-cognitive measure of TSE, and .14 (p=.04) for studies using a non-social cognitive measure of TSE. However, the omnibus test of coefficients was not significant, Q_M (df = 1) = 0.55, p=0.46, indicating that the effect size did not significantly differ based on the two types of TSE scale used, although a trend towards higher effects for social cognitive measures was observed. Average effect sizes for publication format were larger for published studies (r=.19, p=.007) and somewhat smaller for unpublished studies (r = .15, p=.03), but both were significant. The Q_M statistic for this effect was not significant (Q_M (df = 1) = 0.16, p=0.69) indicating that the effect size did not differ significantly based on publication status, consistent with past meta-analyses of teacher self-efficacy (Klassen & Tze, 2014). Table 10 summarizes the four models run: the overall model and the three by domain of effective teaching.

[Insert Table 10 about here]

Analyses of Moderator Effects in the Structured Classroom Management Model

For the 21 effect sizes included in the structured classroom management model, average effect sizes for the type of TES scale were .14 (p=.09) for studies using a social-cognitive measure of TSE, and .08 (p=.41) for studies using a non-social cognitive measure of TSE. The omnibus test of coefficients was not significant, Q_M (df = 1) = 0.30, p=0.58, indicating that the effect size for the association between teacher self-efficacy and structured classroom management did not differ significantly based on the two types of TSE scale used. Average effect sizes for publication status were smaller for published studies (r=.08, p=.39) than for unpublished studies (r=.15, p=.10.), but neither was

significant. The Q_M statistic was also not significant, Q_M (df = 1) = 0.31, p= 0.58, indicating that the effect size did not differ significantly as a function of publication status.

Moderators for the item-level coding were then tested in four separate models. As shown in Tables 11, none of the effect sizes reached statistical significance. However, the addition of the moderator *Problem* (the percentage of items that included an obstacle, challenge or problem), was marginally significant: Q_M (df = 1) = 3.56, 95% CI=[0.00, 0.02], p=0.06, unlike any other models.

[Insert Table 11, Table 12 and Table 13 about here]

Analyses of Moderator Effects in the Supportive Climate Model

Average effect sizes for the type of TSE scale used were .20 (p=.05) for studies using a social-cognitive measure of TSE, and .17 (p=.002) for studies using a non-social cognitive measure of TSE, and both were significant. The omnibus test of coefficients was not significant, Q_M (df = 1) = 0.05, p=0.81, indicating that the effect size for supportive climate did not differ based on the two types of TSE scale used. Average effect sizes were slightly smaller for published studies (r=.16, p=.01) than for unpublished studies (r=.19, p=.008), and both were significantly different from zero. The Q_M statistic was not significant, Q_M (df = 1) = 0.07 p= 0.79, indicating that the effect size did not differ as a function of publication status.

Moderators for the item-level coding were then tested in three separate models. Shown in Table 12, none of the effect sizes were significant, and the estimates were close to 0.

Analyses of Moderator Effects in the Cognitive Activation Model

Average effect sizes for the type of TSE scale used were .17 (p=.04) for studies using a social-cognitive measure of TSE, and .18 (p=.10) for studies using a non-social cognitive measure of TSE. The omnibus test of coefficients was not significant, Q_M (df = 1) = 0.02, p=0.90, indicating that the effect size did not differ based on the two types of TSE scales used. Average effect sizes for publication format were larger for published studies (r=.20, p=.03) and smaller for unpublished studies (r=.15, p=.08). However, the Q_M statistic was not significant, Q_M (df = 1) = 0.18, p=0.67, indicating that the studies' effect sizes did not differ based on publication status.

Moderators for the item-level coding were then tested in three separate models. Shown in Table 13, none of the effect sizes reached statistical significance at .05 level, although the moderator Problem, which assessed percentage of items that included a problem or obstacle for teachers to solve approached significance at p=.059.

Discussion

Compared to prior reviews, the studies included in this meta-analysis focused exclusively on pre-K-12 teachers in core academic subjects (math, science, English and history), and did not include studies that assessed teachers' self-efficacy for aspects of their job other than classroom teaching (e.g., collaboration with colleagues). The meta-analysis detected a statistically significant association between teacher self-efficacy and overall quality of teaching practice (r = .17), but one smaller than reported in a previous meta-analysis of teacher self-efficacy and teaching quality (r = .28; Klassen & Tze, 2014). One possible explanation is the greater number of studies included in this meta-analysis, 39

studies compared to 8. The small effect evident in both studies and the significant amount of unexplained variance suggests that teachers base their sense of efficacy for teaching on other aspects of their teaching besides the quality of their instruction. It may be that contextual factors such as level of professional development support, urbanicity, or percent of students below grade level are moderating variables. Supporting this interpretation, teachers typically evaluate their self-efficacy in comparison to the functioning of other teachers in their schools (Ashton, 1984b), rather than in reference to professional or research-based standards such as those present in the observational measures. Unfortunately, few studies in this meta-analysis reported demographic or contextual information about schools or students. The subsequent chapter explores the meaning of and influences on teachers' efficacy beliefs from the perspective of teachers.

It is still encouraging that a significant association was detected: higher efficacy teachers do tend to exhibit more effective teaching practices, particularly in the domains of supportive classroom climate and cognitive activation. The effect size was not significant in the domain of classroom management, but had a comparable effect size (r=.11 vs. r=[.16,.17]). Interestingly, this was the only model in which a moderating variable, the percentage of items that presented an obstacle, challenge or problem to be solved, approached significance at p=.059. While a larger sample size may provide more robust results, closer examination of the characteristics of the characteristics of TSE scales used in that model is instructive. One TSE scale used in two different studies in the *Structured Classroom Management model* had a large percentage of items that included a problem to be solved specific to classroom management: the 8-item classroom management subscale of the long version of the TSES. Three of the eight items (38%) contained a problem to be

solved, two of which were related to classroom management (The third item about calming a disruptive student was coded as representing the domain of supportive classroom climate because of qualitative findings from Chapter 4). Only one other scale had as high a percentage of items related to solving problems: the Personal Teaching Efficacy subscale of the TES (Gibson & Dembo, 1984). Four of the 9 items on that scale (40%) contained a problem, and 33% of these were a problem related to instruction. However, no studies in the Cognitive Activation model reported an effect size that corresponded with the Personal Teaching Efficacy (PTE) subscale. More commonly, the short version of the TSES, the entire TES, or eclectic scale items were used.

It is also worth noting that sample sizes in the domain-specific models of instruction were small; only 14 effect sizes were detected that had domain matching between the used TSE measure and the observational measure of teaching quality (e.g., TSE for student engagement and teachers' provision of a supportive climate). Such lack of specificity matching limits the ability to draw precise conclusions about the nature of the relationship between teachers' domain-specific beliefs and their domain-specific behaviors. While the percentage of items per scale that corresponded with the domain of teaching (in the domain-specific models) was not a significant moderator, it may be that although the items matched the general domain of teaching, they did not represent the breadth or depth of practices captured by the observational tool used in the study. For example, teacher self-efficacy measures tend to focus more singularly on efficacy for management of disruptive behavior and discipline while the Classroom Observation System focuses on many other indicators of a well-managed and structured classroom. This suggests that the conceptualization and measurement of teacher self-efficacy for

classroom management may need to be broadened to include other important aspects of classroom management beyond behavior management such as clarity of instruction and efficient use of resources. Doing so may improve the predictive validity of this domain for teacher practices that matter for student learning.

Similarly, it is unclear whether the observational tools in each study adequately captured the breadth of the domains of effective instruction. Recall that if 30% of items on a scale could be categorized in one domain, then the scale was categorized as representative of that domain. This does not imply that the breadth and depth of the observational tool matches the domain as described. While it is outside the scope of this dissertation to assess the degree of match between the observer measures used and the measure of TSE used, this is an important area for future investigation for which the coding system and findings have laid the groundwork.

None of the other tested moderator variables (type of TSE measure or publication status) were significant in any of the models. While other studies with larger sample sizes have detected effects of publication status (Hulleman et al., 2010) and of type of measure used (Eells, 2012), Klassen & Tze (2014) did not detect evidence of publication bias in their meta-analysis of TSE and teaching performance. Average effect sizes did differ based on these variables, though these differences did not reach statistical significance and were not always in the same direction. For example, in the model for cognitive activation, published studies (k=7) demonstrated a higher mean effect size (.20) than unpublished studies (k=11, r=.14). This trend is consistent with previous meta-analytic findings, yet the small sample size may have limited the available statistical power needed to detect a

statistically significant effect with these categorical moderators. Therefore, further research on the effects and implications of teachers' self-efficacy is needed.

Finally, nearly half of all studies used measures of TSE that have demonstrated psychometric problems and are generally acknowledged to be flawed conceptually. However, use of these measures did not seem to impact the strength of the relationship between TSE and teaching practice significantly, assessed through moderator analyses. Nevertheless, there was a trend in the data toward larger effects for scales using social-cognitive theory and an effect of including problems in the TSE item. An open question, therefore, is not only the practical implications of TSE (Klassen et al., 2014), but also the practical implications of the social-cognitive shift in conceptualization and measurement of TSE that ushered in a new era of TSE research (Klassen, Tze, Betts, & Gordon, 2011). Using teacher self-efficacy scales with stronger psychometric properties and greater theoretical alignment, one might anticipate detecting stronger effects on teaching behavior, as self-efficacy is typically a strong predictor of behavior (Bandura et al., 1996). However, teaching is a dynamic, interpersonal and highly contextual profession: behavior is influenced by a number of factors beyond personal beliefs, such as available resources in the environment, pedagogical focus of the school, knowledge and skill in teaching, professional development, and characteristics of the students themselves. It may be that the meaning of teachers' self-efficacy beliefs is dependent on such factors or changes depending on how TSE assessed and how teachers judge their sense of efficacy, the focus of the next chapter.

Table 5.

Indicators of effective classroom teaching by domain

	Structured Classroom Management	Supportive Classroom Climate	Cognitive Demand
Specific Indicators	*A. Opportunity to learn (classroom organization)* + momentum, brisk pacing of lessons + effective transitions + routines and procedures for learning + high student engagement (behavioral) *B. Proactive behavior management* + nipping potential problems in the bud + consistency of rule enforcement + providing clues to situational expectations + effective treatments of interruptions *C. Instructional coherence* + clear instructional goals + curriculum alignment + coherent content (coherent explanations, clear directions, vivid descriptions) + goal-oriented assessments + ongoing formative assessment	*A. Positive teacher-student interaction* + teacher enthusiasm, warmth (e.g. smiling) + praise + teacher knows students as individuals and as learners + responsiveness to student input (e.g., integration of student ideas) + appropriate expectations + orientation to students' personal needs + process emphasized as much as product (mastery vs. performance goals) + responsive to individual differences: seen as opportunities for learning rather than obstacles to be overcome + instruction in self-regulation strategies, if needed + support for student non-cognitive outcomes of self-efficacy, agency, and self-monitoring *B. Positive student-student interaction* + respectful interactions + mutual respect + cooperation + high student engagement (affective) *C. Cooperative activity structure* + cooperative learning + heterogeneous grouping + flexible grouping arrangements + adaptation of represented materials to the characteristics of the context/students (not individualized instruction in which there are different learning goals for different students)	*A. Thoughtful discourse* + higher level cognitive questions + teachers probe for explanations + constructive feedback + mastery assessed through extended responses (orally or in writing) + public questions (response opportunities) + pose questions with multiple answers + students discuss ideas by talking (construction of knowledge) + encouraging students to relate class work to their own experiences + careful attention and diagnosis of students' knowledge + scaffolding students' ideas and task involvement + high student engagement (cognitive) *B. Metacognitive strategies* + reflection + prompt students to explain thinking and how they arrived at conclusions + strategy teaching (models and instructs learning, problem solving) + debriefing learning activities + establishing learning orientations (clarifying intended outcomes and cuing desired learning strategies) *C. Problem solving* + Inquiry models & authentic activities + practice/application + interpretation of findings + opportunities for extended writing projects

Table 6.

TSE measures used in studies

	# studies	Classroom management	Student engagement	Instructional strategies
TSES (current)	k=11	x	x	x
TES (past)	k=12			
modified TES	k=3			
Bandura's TSES	k=3	x	x	x
STEBI (past)	k=1			
S&S	k=1	x	x	x
CMDS	k=1	x		

Table 7.

Observational measures used in studies

	# studies	Structured classroom management	Supportive Climate	Cognitive activation
CLASS	k=7	x	x	x
FFT / THOR	k=6	x	x	x
RTOP	k=3		x	x
OCIW	k=2	x		x
CPI	k=1		x	x
ECCOM	k=1	x	x	x
on-task behavior	k=2	x		
unique to study	k=20			

Table 8.

Domain matching in coded studies

	Teacher self-efficacy for management	Teacher self-efficacy for engagement	Teacher self-efficacy for instruction	Multi-dimensional social cognitive scales (e.g. TSES)	Personal Teaching Efficacy (PTE)**	Multi-dimensional, non-social cognitive scales (e.g., TES)	Total studies (effect sizes)
Structured management	***k =5****	*k=1*	*k=1*	k=6	k=4	k=10	k=21 (27)
Supportive climate	k =3	***k=4***	k=2	k=5	k=0	k=5	k=14 (19)
Cognitive activation/ demand	k=2	k=2	***k=5***	k=8	k=1	k=8	k=18 (35)
Multi-dimensional measure of teaching effectiveness	k=5	k=3	k=4	***k=15***	k=5	k=14	k=33

* k within a cell represent independent effect sizes; each row may include dependent effect sizes because some contributed more than one effect size (e.g., reported correlations for subscales and entire measure or included multiple measures of teacher self-efficacy or teaching effectiveness)

** one study assessing PTE had insufficient information to code for domain of teaching effectiveness assessed

Table 9.

All studies meeting eligibility criteria

Author	Year	Publication status	n	Years Teaching Average	Teacher self-efficacy measure	Observational measure of teaching	Structured Management domain	Supportive Climate domain	Cognitive Activation domain	Number of effect sizes	Effect size (Pearson r)
Almog	2007	article	33	n/r	TES	unique		x		1	.30
Anderson (1.1)	1988	article	24	n/r	TES	OCIW			x	1	0
Anderson (1.2)	1988	article	24	n/r	TES	OCIW			x	1	.29
Anglin-Bodrug	2005	dissertation	32	0	TSES	unique		x		1	-.13
Ashley	2009	dissertation	8	4	TSES	unique	x		x	5	.45
Bozack	2006	conference	10	1	TES-like	FFT/THOR	x		x	3	.90
Bozack 2.1	2007	conference	14	1	TES-like	FFT/THOR	x		x	3	-.74
Bozack 2.2	2007	conference	14	1	TES-like	FFT/THOR	x		x	3	.29
Bozack 2.3	2007	conference	14	1	TES-like	FFT/THOR	x		x	3	.16
Bozack	2008	dissertation	295	1	TES-like	FFT/THOR	x		x	3	-.20
Brown	2003	dissertation	20	8	TSES	unique		x	x	1	.24
Cantrell	2008	article	22	11	modified TES	unique			x	2	.49
Collmann	2012	dissertation	317	14	TSES	CLASS	x	x	x		
Emmer	1991	article	30	0	TES-like	unique	x			3	-.05
Guo	2010	article	67	15	Bandura TSES	CLASS			x	3	.17
Guo	2012	article	896	15	Bandura TSES	CLASS	x	x		5	.08
Heneman	2006	report	180	11	TSES	FFT/THOR	x	x	x	9	.35
Holzberger	2013	article	155	22	S&S	unique	x	x	x	3	.30
Jablonski	1997	dissertation	15	0	unique	unique	x	x		6	-.03

Jamil	2012	article	509	0	TSES	CLASS	x	x	x	3	0
Jeck	2010	dissertation	38	13	TES	unique			x	2	.11
Justice	2008	article	135	15	Bandura TSES	unique	x		x	4	.11
Kopcha	2011	article	17	0	TSES	unique			x	2	.09
Lardy	2011	dissertation	36	19	STEBI (PSTE)	RTOP		x	x	3	.33
Lerkanen	2012	article	49	15	TES	ECCOM		x	x	1	.12
Moore	2008	dissertation	89	0	TSES	unique		x	x	3	.29
Nissim	2003	dissertation	16	7	TES	unique		x	x	6	.56
O'Connor	1998	dissertation	50	16	TES	unique		x		1	-.05
Ortega	2009	dissertation	32	8	TES	unique		x		4	.07
Pakarinen	2010	article	48	8	TES	CLASS	x	x	x	3	.16
Reinke	2013	article	33	13	TSES	unique	x	x		4	.27
Rohs	2007	dissertation	61	13	TSES	CPI		x		12	.06
Saklofske	1988	report	65	0	TES	unique	x			3	.24
Spilt	2012	article	32	13	TES	CLASS	x	x		2	.14
Stanovich	1998	article	33	10	TES-like	unique		x		1	-.12
Stein	1988	article	14	13	TES-like	unique		x	x	1	-.34
Teel	2003	dissertation	102	10	TES	unique	x	x		6	.05
Temiz	2013	article	101	0	TSES	RTOP		x	x	15	.78
Tracz	1986	report	14	n/r	TES	on-task behavior	x			2	.18
Warren	2009	dissertation	4	2	TSES	RTOP		x	x	4	-.86*

TES=Teacher Efficacy Scale; TSES=Teacher Sense of Efficacy Scale; Bandura TSES= Bandura's unpublished Teacher Sense of Efficacy scale, used in NICCHD studies; STEBI (PSTE)=Science Teaching Efficacy Beliefs Instrument (PSTE scale); S&S=Schwarzer & Schmidt Teacher Efficacy Scale

OCIW=Our Class and Its Work; FFT/THOR=Framework for Teaching or Tsang-Hester Observation Rubric; CLASS = Classroom Assessment Scoring System; RTOP = Reformed Teaching Observation Protocol; ECCOM=Early Childhood Classroom Observation Measure; CPI=Classroom Practices Inventory

* excluded from statistical analyses; included in all descriptive analyses

Table 10.

Summary of effect sizes for full model and by domains of teaching practice

	k	r	SE	z	95% CI
Full model (all teaching domains)	39	.17***	0.05	3.51	[0.08, 0.27]
Structured management	21	.11	0.06	1.85	[0.00, 0.22]
Supportive Climate	14	.17***	0.05	3.69	[0.08, 0.25]
Cognitive Activation	18	.17**	.06	2.77	[0.05, 0.29]

***$p < .001$ **$p < .01$

Table 11.

Summary of moderators for Structured Management model (k=21)

	estimate	SE	Z	95% CI
Percent of items assessing management	.00	0.00	.13	[-0.00, 0.01]
Percent of items with future orientation	.00	.00	.72	[0.00, 0.00]
Percent of items with problem of practice	.01	0.00	1.88	[0.00, 0.02]

Table 12.

Summary of moderators for Supportive Climate model (k=14)

	estimate	SE	Z	95% CI
Percent of items assessing climate	.00	0.00	-.63	[-0.01, 0.00]
Percent of items with future orientation	.00	.00	-.20	[0.00, 0.00]
Percent of items with problem of practice	.01	0.01	.42	[-0.02, 0.02]

Table 13.

Summary of moderators for Cognitive Activation model (k=18)

	estimate	SE	Z	95% CI
Percent of items assessing climate	.00	0.00	-.13	[-0.00, 0.00]
Percent of items with future orientation	.00	.00	-.03	[-0.00, 0.00]
Percent of items with problem of practice	.00	0.01	.63	[-0.01, 0.02]

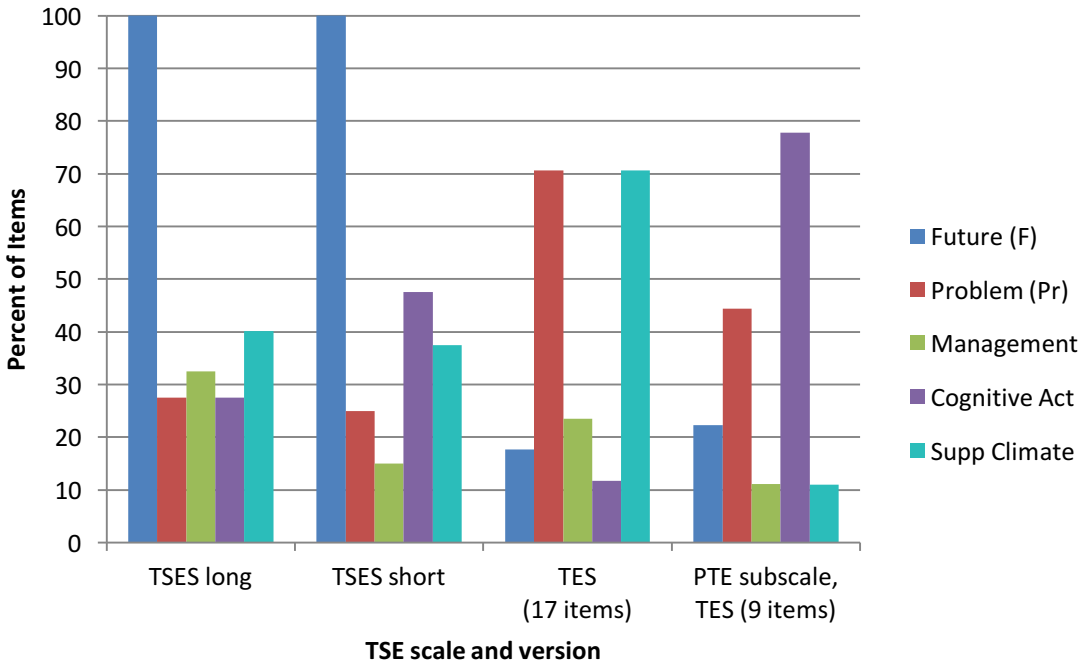

Figure 5. Moderator coding by teacher self-efficacy scale

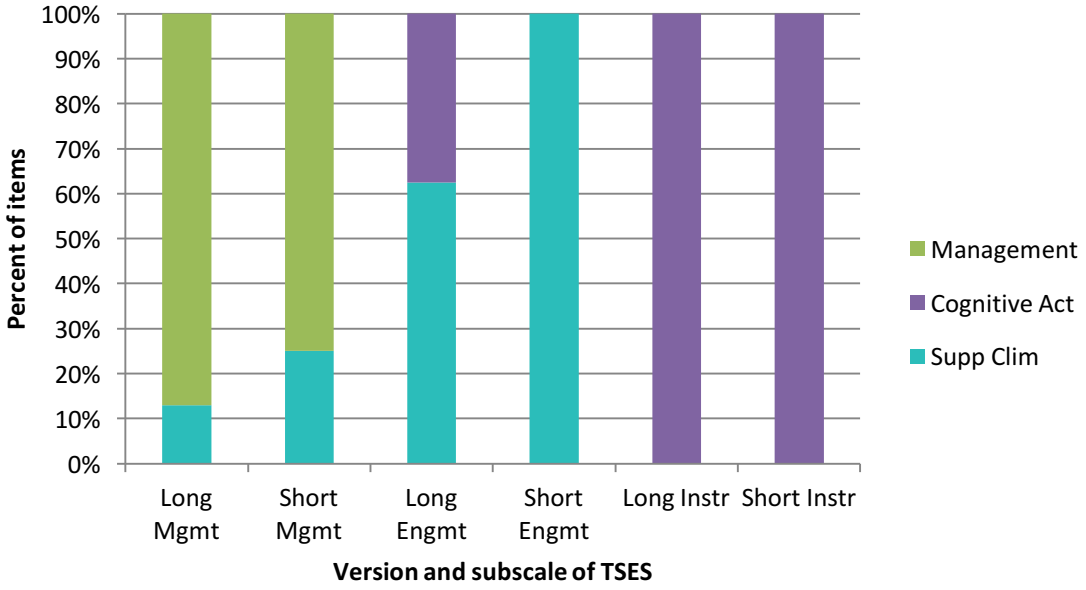

Figure 6. Domain coding by version and subscale of TSES

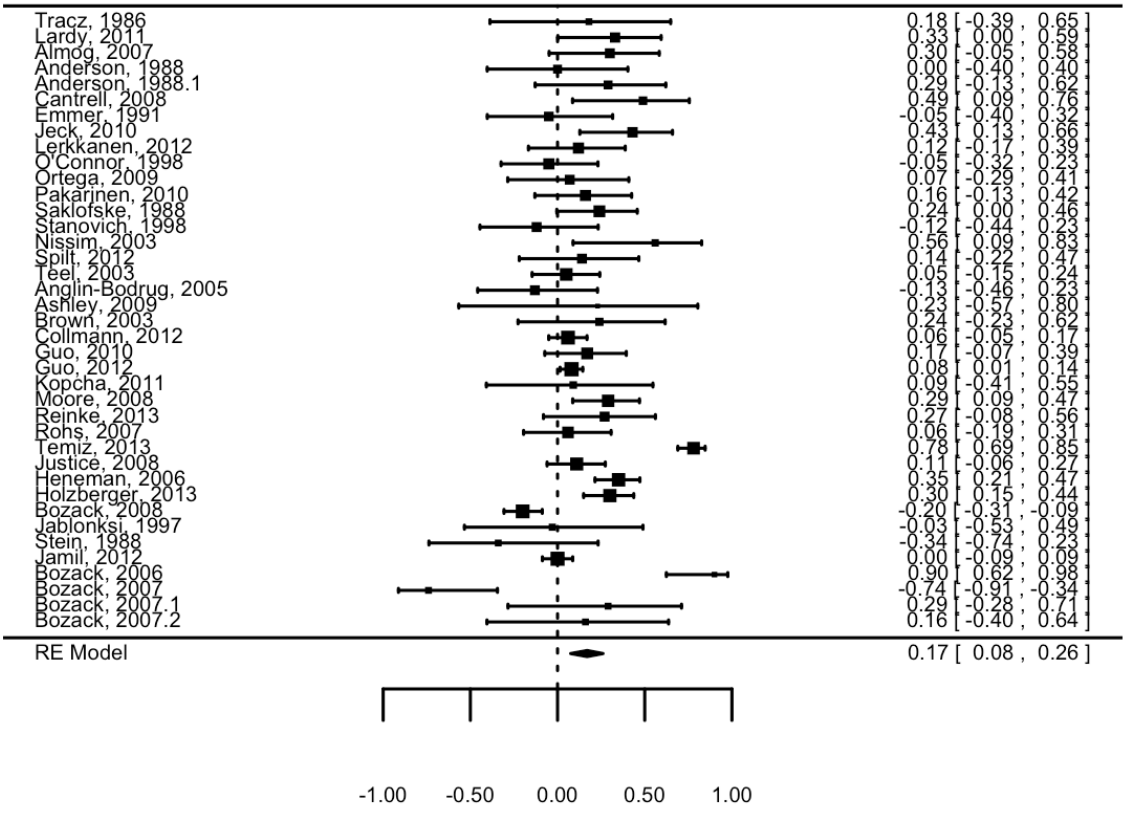

Figure 7. Synthesis forest plot for included studies

References

Allinder, R. M. (1995). An examination of the relationship between teacher efficacy and curriculum-based measurement and student achievement. *Remedial and Special Education*, *16*(4), 247–254. http://doi.org/10.1177/074193259501600408

Armor, D., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., … Zellman, G. (1976). *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools.* Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED130243

Ashton, P. T. (1984a). Teacher Efficacy: A Motivational Paradigm for Effective Teacher Education. *Journal of Teacher Education*, *35*(5), 28–32.

Ashton, P. T. (1984b). Teachers' Sense of Efficacy: A Self- or Norm-Referenced Construct? *Florida Journal of Educational Research*, *26*(1), 29–41.

Ashton, P. T., & Webb, R. B. (1986). *Making a difference: teachers' sense of efficacy and student achievement*. New York: Longman. Retrieved from http://hdl.handle.net/2027/[u]: mdp.39015010602426

Balam, E. M. (2006). *Professor's Teaching Effectiveness in relation to self-efficacy belifs and perceptions of student rating myths*. Auburn University.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY US: W H Freeman/Times Books/ Henry Holt & Co.

Bandura, A. (2006). Guide for Constructing Self-Efficacy Scales. In *Self-Efficacy Beliefs of Adolescents* (pp. 307–337). Information Age Publishing.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development*, *67*(3), 1206–1222. http://doi.org/10.2307/1131888

Barnes, G. L. V. (1998, April). *A comparison of self-efficacy and teaching effectiveness in preservice string teachers*. Ohio State University, US.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., … Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180. http://doi.org/10.3102/0002831209345157

Blazevski, J. L. (2006). *Teacher efficacy for supporting student motivation*. ProQuest Information & Learning, US.

Bong, M. (2006). Asking the right question: How confident are you that you could successfully perform these tasks? In *Self-Efficacy Beliefs of Adolescents* (pp. 287–305). Information Age Publishing.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40. http://doi.org/10.1023/A:1021302408382

Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, *40*(2), 119–156.

Brophy, J. E. (2006). Observational Research on Generic Aspects of Classroom Teaching. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 755–780). Mahwah, NJ: Erlbaum.

Brophy, J. E., & Good, T. L. (1986). Teacher behaviors and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). New York: Macmillan.

Brouwers, A., Tomic, W., & Stijnen, S. (2002). A confirmatory factor analysis of scores on the teacher efficacy scale. *Swiss Journal of Psychology*, *61*(4), 211–219. http://doi.org/10.1024//1421-0185.61.4.211

Chang, M.-L. (2013). Toward a theoretical model to understand teacher emotions and teacher burnout in the context of student misbehavior: Appraisal, regulation and coping. *Motivation and Emotion*, *37*(4), 799–817. http://doi.org/10.1007/s11031-012-9335-0

Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education*, *28*(3), 258–277. http://doi.org/10.2307/749781

Cohen, S. A., Hyman, J. S., Ashcroft, L., & Loveless, D. (1989). Comparing effects of metacognition, learning styles, and human attributes with alignment. In *Annual meeting of the American Educational Research Association*. San Francisco, CA.

Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education*, *60*(4), 323–337. http://doi.org/10.1080/00220973.1992.9943869

Corno, L. (2008). On Teaching Adaptively. *Educational Psychologist*, *43*(3), 161–173. http://doi.org/10.1080/00461520802178466

Darling-Hammond, L., & Bransford, J. (2005). *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass, An Imprint of Wiley.

Dolan, A. L., & McCaslin, M. (2008). Student perceptions of teacher support. *Teachers College Record*, *110*(11), 2423–2437.

Eccles, J. S., & Roeser, R. W. (1998). School and Community Influences on Human Development. In M. H. Borenstein & M. E. Lamb (Eds.), *Developmental Psychology: An advanced textbook* (4th ed., pp. 503–554). Mahwah, NJ: Erlbaum.

Eells, R. J. (2012). *Meta-analysis of the relationship between collective teacher efficacy and student achievement*. ProQuest Information & Learning, US.

Emmer, E. T., & Hickman, J. (1991). Teacher Efficacy in Classroom Management and Discipline. *Educational and Psychological Measurement*, *51*(3), 755 –765. http://doi.org/10.1177/0013164491513027

Enochs, L. G., Scharmann, L. C., & Riggs, I. M. (1995). The relationship of pupil control to preservice elementary science teacher self-efficacy and outcome expectancy. *Science Education*, *79*(1), 63–75. http://doi.org/10.1002/sce.3730790105

Evertson, C. M., Anderson, C. W., Anderson, L. M., & Brophy, J. E. (1980). Relationships Between Classroom Behaviors and Student Outcomes in Junior High Mathematics and English Classes. *American Educational Research Journal*, *17*(1), 43–60. http://doi.org/10.3102/00028312017001043

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, *76*(4), 569–582. http://doi.org/10.1037/0022-0663.76.4.569

Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, *31*(3), 627–643. http://doi.org/10.2307/1163230

Holden, H., & Rada, R. (2011). Understanding the Influence of Perceived Usability and Technology Self-Efficacy on Teachers' Technology Acceptance. *Journal of Research on Technology in Education*, *43*(4), 343–367.

Holzberger, D., Philipp, A., & Kunter, M. (2013). How Teachers' Self-Efficacy Is Related to Instructional Quality: A Longitudinal Analysis. *Journal of Educational Psychology*. http://doi.org/10.1037/a0032198

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, *136*(3), 422–449. http://doi.org/10.1037/a0018947

Kazemi, E. (1998). Discourse That Promotes Conceptual Understanding. *Teaching Children Mathematics*, *4*(7), 410–14.

Klassen, R. M., Durksen, T. L., & Tze, V. M. C. (2014). Teachers' self-efficacy beliefs: Ready to move from theory to practice? In P. W. Richardson, S. A. Karabenick, & H. M. G. Watt (Eds.), *Teacher motivation: Implications for theory and practice* (pp. 100–115). U.S.: Routledge/Taylor & Francis.

Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, *12*, 59–76. http://doi.org/10.1016/j.edurev.2014.06.001

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher Efficacy Research 1998–2009: Signs of Progress or Unfulfilled Promise? *Educational Psychology Review*, *23*(1), 21–43. http://doi.org/10.1007/s10648-010-9141-8

Klieme, E. (2012, August). *Qualities and effects of teaching: Towards a conceptual theory of teaching.* Presented at the EARLI SIG meeting, Educational Effectiveness, Zurich, Switzerland.

Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York: Holt, Rinehart and Winston. Retrieved from http://hdl.handle.net/2027/[u]: mdp.39015004798677

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, Calif.: Sage Publications. Retrieved from http://hdl.handle.net/2027/[u]: mdp.39015049694659

McWilliams, M. A., Nier, J. A., & Singer, J. A. (2013). The implicit self and the specificity-matching principle: Implicit self-concept predicts domain-specific outcomes. *Personality and Individual Differences*, *54*(4), 474–478. http://doi.org/10.1016/j.paid.2012.09.014

Newmann, F. M. (1996). *Authentic achievement: restructuring schools for intellectual quality*. San Francisco: Jossey-Bass Publishers. Retrieved from http://hdl.handle.net/2027/[u]: mdp.39015038157007

O'Neill, S. C., & Stephenson, J. (2011). The measurement of classroom management self-efficacy: A review of measurement instrument development and influences. *Educational Psychology*, *31*(3), 261–299. http://doi.org/10.1080/01443410.2010.545344

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity. *Educational Researcher*, *38*(2), 109–119. http://doi.org/10.3102/0013189X09332374

Pianta, R. C., La Paro, K. M., Payne, C., Cox, M. J., & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school characteristics and child outcomes. *The Elementary School Journal*, *102*(3), 225–238. http://doi.org/10.1086/499701

Preiss, R. W., Gayle, B. M., & Allen, M. (2006). Test Anxiety, Academic Self-Efficacy, and Study Skills: A Meta-Analytic Review. In B. M. Gayle, R. W. Preiss, N. Burrell, M.

Allen, B. M. (Ed) Gayle, R. W. (Ed) Preiss, … M. (Ed) Allen (Eds.), *Classroom communication and instructional processes: Advances through meta-analysis.* (pp. 99–111). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Rohrkemper, M., & Corno, L. (1988). Success and failure on classroom tasks: Adaptive learning and classroom teaching. *The Elementary School Journal*, *88*(3), 297–312. http://doi.org/10.1086/461540

Ryan, A. M., Kuusinen, C. M., & Bedoya-Skoog, A. (2015). Managing peer relations: A dimension of teacher self-efficacy that varies between elementary and middle school teachers and is associated with observed classroom quality. *Contemporary Educational Psychology*, *41*, 147–156. http://doi.org/10.1016/j.cedpsych.2015.01.002

Skaalvik, E. M., & Skaalvik, S. (2007a). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*(3), 611–625. http://doi.org/10.1037/0022-0663.99.3.611

Skaalvik, E. M., & Skaalvik, S. (2007b). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*(3), 611–625. http://doi.org/10.1037/0022-0663.99.3.611

Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, *26*(4), 1059–1069. http://doi.org/10.1016/j.tate.2009.11.001

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. http://doi.org/10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., & Johnson, D. (2011). Exploring literacy teachers' self-efficacy beliefs: Potential sources at play. *Teaching and Teacher Education*, *27*(4), 751–761. http://doi.org/10.1016/j.tate.2010.12.005

Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, *73*(1), 287–301. http://doi.org/10.1111/1467-8624.00406

Williams, W. M., Blythe, T., White, N., Li, J., Gardner, H., & Sternberg, R. J. (2002). Practical intelligence for school: Developing metacognitive sources of achievement in adolescence. *Developmental Review*, *22*(2), 162–210. http://doi.org/10.1006/drev.2002.0544

Woolfolk, A. E., & Hoy, W. K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology*, *82*(1), 81–91. http://doi.org/10.1037/0022-0663.82.1.81

CHAPTER 4

THE MEANING OF TEACHERS' SELF-EFFICACY BELIEFS: PERSPECTIVES OF MIDDLE AND HIGH SCHOOL TEACHERS

Introduction

Substantial evidence indicates that teachers' sense of efficacy—teachers' confidence in their capability to execute courses of action necessary to produce desired educational outcomes—has critical implications for teachers' job satisfaction, intent to stay in the profession, and burnout (Kim & Kim, 2010; Moè, Pazzaglia, & Ronconi, 2010; Skaalvik & Skaalvik, 2007; Wang, Hall, & Rahimi, 2015). Furthermore, a recent meta-analysis found a positive, though relatively small, correlation between teacher self-efficacy and student achievement when assessing effect sizes across samples consisting of novice and experienced teachers in both academic (e.g., English) and non-academic subjects (e.g., music education) (Klassen & Tze, 2014). Less is known however about potential differences in how teachers conceptualize their self-efficacy for teaching, what evidence they use to judge their degree of self-efficacy, and the implications of teachers' self-efficacy for teachers' instructional behaviors (Klassen, Durksen, & Tze, 2014; Kleinsasser, 2014). Indeed, recent reviews of the literature suggest that despite prolific investigations, teacher efficacy research is not quite ready to move from theory to practice. There is a need to further examine the links between teacher efficacy and educational outcomes, and the vast majority of available evidence is correlational, which provides only limited insight into the role of teachers' self-efficacy in the instructional process (Klassen et al., 2014).

Furthermore, researchers have pointed out that the complex, multi-faceted nature of teacher self-efficacy presents unique challenges for developing reliable and nuanced measures that are valid across the multitude of contexts in which teachers work (Kleinsasser, 2014; Labone, 2004; Wheatley, 2005). Qualitative research provides the opportunity to illuminate measurement problems in the study of teacher self-efficacy as well as further develop theory about how teachers form their self-efficacy beliefs.

The present qualitative study aims to expand upon prior evidence by focusing on the processes by which teachers rate their sense of self-efficacy, including the extent and manner in which work experience and student outcomes play a role in teachers' self-judgments. Understanding the links between teachers' self-efficacy and educational outcomes requires not only systematic quantitative analyses, but also qualitative analyses into teachers' thinking about self-efficacy and its implications for the instructional process. Specifically, think-aloud methodology and cognitive interviewing techniques (Karabenick et al., 2007) were utilized in the present study to examine teachers' responses to the Teachers' Sense of Efficacy Scale (Tschannen-Moran & Hoy, 2001), which is the most commonly used assessment of teacher self-efficacy (Klassen, Tze, Betts, & Gordon, 2011). This methodological approach provides access to teachers' thinking while they reflect upon self-efficacy indicators that are typically used in the scientific community (i.e., the TSES scale items), and to the ways that teachers evaluate their sense of efficacy in relation to these indicators. It builds on past research that has demonstrated the strong psychometric properties of the TSES across cultural contexts (Duffin, French, & Patrick, 2012; Tsigilis, Koustelios, & Grammatikopoulos, 2010) and addresses the need for more qualitative studies of teacher self-efficacy, particularly regarding the interpretation of self-

efficacy beliefs (Coladarci & Breton, 1991; Klassen et al., 2011; Kleinsasser, 2014; Tschannen-Moran & Hoy, 2001; Wheatley, 2005). Whereas quantitative research produces only numeric scores in response to scale items (e.g., a "1" or a "2" may indicate low levels of efficacy for a particular outcome), however, the present study also collected qualitative evidence to aid the interpretation of these numeric scores (e.g., why teachers rate themselves as a "1" or a "2" and how they interpret such numeric scores).

Specifically, I examined: (a) the meaning of teacher self-efficacy and views of the teaching task as a function of specific features of the self-efficacy item (e.g., whether or not an explicit reference to student outcomes is included) and background characteristics (e.g., years of work experience, professional training); (b) potential qualitative differences in teachers' self-efficacy judgments as a function of such background characteristics, since prior research suggests that these factors have quantifiable effects on teachers' self-judgments; and (c) the overall content validity of the TSES items and subscales from the perspective of teachers (e.g., the extent to which teachers' judgments seem to be captured by the theoretical framework underlying the TSES, and potential sources of ambiguity). In the following section I review the theoretical foundation for the TSES and expand upon each of these research objectives.

Literature Review

Teacher self-efficacy research and the need for qualitative analyses

Teacher self-efficacy research is based in social cognitive theory and assesses future-oriented beliefs about whether teachers can execute courses of action at a desired level of competence (Bandura, 1997). These beliefs are strong indicators of behavior;

teachers are unlikely to perform tasks that they believe they lack the skills or resources to accomplish successfully (Cantrell & Callaway, 2008; Ross, McKeiver, & Hogaboam-Gray, 1997). The majority of empirical evidence shows that teacher self-efficacy is positively related to important teacher outcomes such as job satisfaction and intent to stay in the profession, and negatively related to burnout (Brouwers & Tomic, 2000; Skaalvik & Skaalvik, 2007, 2010)

The most impactful sources of teachers' sense of efficacy are their mastery experiences, or their subjectively perceived instances of success or failure in teaching (Bandura, 1997). Such efficacy experiences provide the foundation for two cognitive processes that Tschannen-Moran and colleagues (Tschannen-Moran & Hoy, 2001; Tschannen-Moran, Hoy, & Hoy, 1998) propose teachers engage in when prompted by self-efficacy indicators such as survey questions: the analysis of the teaching task and the judgment of their own teaching competence. In the analysis of the teaching task, teachers first ask themselves, "What outcomes do I seek, that is, what is success in this teaching task?" (Tschannen-Moran et al., 1998, p. 232), a means-end analysis (Skinner, 1996). Teachers then reflect upon their capabilities to master this task—the judgment of their own competence. While much research has focused on the implications of quantitative differences in teachers' self-efficacy judgments, comparatively little research has focused on the meaning of these outcome-oriented judgments from the teachers' perspectives (see Tschannen-Moran & McMaster, 2009 for an exception). Such an investigation is critical to understand to what extent and in what situations teachers' conceptualization of success is dependent on students' success, a critical association for investigating how responsive teachers are to students' needs, an important characteristic of effective teaching (Corno,

2008; Klieme, 2012; Pianta, Belsky, Vandergrift, Houts, & Morrison, 2008) Accordingly, the present study addresses how teachers analyze the teaching task (i.e., define success in the teaching task) and its effect on subsequent judgments of competence.

The role of students

Studies that have used the TSES (TSES; Tschannen-Moran & Hoy, 2001), a scale widely accepted to be both congruent with self-efficacy theory and psychometrically valid across cultural contexts, have found no relationship (Domsch, 2009; Donald, 2009), or a modest relationship with student achievement (Caprara, Barbaranelli, Steca, & Malone, 2006; Ross, Hogaboam-Gray, & Hannay, 2001). Researchers' struggles to find strong associations between the TSES and student outcomes such as achievement pose an important question: To what extent are student outcomes such as achievement and engagement relevant for teachers' self-judgments, and if student outcomes are not strongly taken into consideration, what is the foundation for teachers' self-judgments?

Some theorists have suggested that student outcomes such as students' increased understanding or motivation should be viewed as *indicators* of teaching competence (D. K. Cohen, 2011; Grossman et al., 2007; Lampert, Boerst, & Graziani, 2011). In this way, student outcomes would be an inherent part of self-efficacy judgments, the "desired level of competence" to which teachers strive. On the other hand, student outcomes could also be considered the expected outcomes of successful teaching (Bandura, 1997; Fenstermacher & Richardson, 2005). In other words, teaching would be the task for which teachers may feel more or less efficacious, whereas student learning or engagement would be an outcome expectancy—i.e., the expected consequence of performing the teaching task

at some desired level of competence (Bandura, 1997). According to this conceptualization, teachers' self-efficacy judgments would be solely based on teachers' confidence in their ability to perform certain teaching behaviors, which then may or may not lead to student learning. The implications of this distinction for teachers' self-judgments have not been examined systematically. However, it seems plausible to assume that teachers who rate their efficacy in relation to their own teaching behaviors (e.g., lesson plan preparation or knowledge of teaching methods) rather than student outcomes (e.g., how much students learn from their lessons) would be less vulnerable to student failure and may generally rate themselves more positively, since they would be less likely to subjectively experience failure. Accordingly, an important objective of the present study was to examine teachers' own definitions of teaching success (in reference to own teaching behaviors versus actual student outcomes), as well as potential implications of this distinction for their self-perceptions and perceptions of students.

The role of teachers' experience

Researchers have suggested that novice and experienced teachers' self-reported efficacy may mean different things and lead to different interpretations of responses to survey items (Henson, 2001; Tschannen-Moran et al., 1998; Wheatley, 2005). From a cognitive perspective, novice teachers' thinking is typically characterized as focused on retrieval of pedagogical strategies that have not yet become automatized. In complex and cognitively demanding situations such as teaching, novice teachers may be more likely to focus on their own behaviors (How should I teach in this situation?) rather than student behaviors (How are students responding to my teaching?), as a means of reducing the

complexity of the classroom (Feldon, 2007a). In such an evaluation, their perceived level of pedagogical knowledge—knowledge of teaching strategies—might be a more salient factor in their efficacy judgments. In other words, capability to generate a lengthy inventory of strategies may lead to high sense of efficacy even if those strategies have not yet been successful in bringing about desired student outcomes. On the contrary, other scholars have suggested that novices may base their efficacy judgments on what they believe is possible to accomplish (e.g., I'm confident I can learn to do this) rather than what they currently know how to do (Tschannen-Moran et al., 1998). Accordingly, the present study will investigate the extent to which novices are consistent in how they judge their self-efficacy, particularly across different domains of teaching tasks (e.g., instruction, student motivation, and classroom management), and to different types of survey items (i.e., outcome or strategy-based indicators).

When novices do consider student responses to instructional strategies, they may be more likely to focus on non-cognitive outcomes of instruction (Did they enjoy it?) rather than factors such as student learning (Did they understand?) compared to more experienced teachers (Mulholland & Wallace, 2001). Novices must typically be coached by more experienced colleagues to consider student learning as an outcome of successful instruction and as relevant criteria by which to judge their efficacy (Gabriele & Joram, 2007). This consideration of non-cognitive outcomes of instruction may contribute to problems in discrimination between efficacy subscales of instruction and student engagement for novice teachers (Duffin et al., 2012), potentially limiting the predictive validity of novices' instructional efficacy for student learning outcomes.

Moreover, failure to achieve student learning or enjoyment outcomes may differentially affect novice and experience teachers' sense of efficacy. Experts have greater problem solving abilities because of their greater ability to diagnose patterns and circumvent common problems (Borko & Livingston, 1989; Feldon, 2007b). When confronted with difficulties or failure to achieve student goals, more experienced teachers may focus their sense of efficacy on whether they can understand *why* students failed to understand, such as through eliciting information about student thinking and diagnosing potential misconceptions in their content areas (Ball & Forzani, 2010; Coffey, Hammer, & Levin, 2009; Jacobs, Lamb, & Philipp, 2010; Kazemi, 1998). Problem-solving efficacy is typically lowest among beginning teachers (Chang, 2013), suggesting that more novice teachers' efficacy judgments may be more impacted by failure to achieve desired outcomes of instruction than the efficacy judgments of their more experienced colleagues. However, more experienced teachers may also have higher standards for success compared to their novice colleagues. This may lead to a reference bias (Duckworth & Yeager, 2015) in which more experienced teachers rate themselves more harshly and feel less successful yet might objectively be perceived as more effective.

In summary, much is unknown about the meaning of teachers' efficacy beliefs in relation to student outcomes, years of experience, and past learning experiences. Inconsistency and inattention to the definition of the teaching task and its relation to outcome expectancies continue to limit the conclusions one can draw about the meaning of teacher self-efficacy beliefs in relation to student outcomes.

The present study

The present study contributes to this debate by acknowledging teachers' perspectives (how they define the teaching task) and corresponding implications (do they rate themselves differently as a consequence of how they define the teaching task; do they think about students or the task differently)? By addressing these questions, the present study endeavored to provide greater insight into the construct and predictive validity of teacher self-efficacy scales.

Specifically, this involved capturing novice and experienced teachers' interpretations of their efficacy beliefs in response to items from the short version of the TSES (Tschannen-Moran & Hoy, 2001). The approach draws on cognitive interviewing (Karabenick et al., 2007) and protocol analysis (Ericsson & Simon, 1993). Only middle and high school teachers were included in the analyses, so that all participants were teaching similar age groups. In addition, focus on secondary teachers is important because they tend to report lower levels of self-efficacy relative to elementary teachers, the reasons for which are only partially understood (Guskey, 1987; Ryan, Kuusinen, & Bedoya-Skoog, 2015). Because self-efficacy beliefs are highly contextual (Bandura, 1997; Raudenbush, Rowan, & Cheong, 1992), I also attended to the influence of context (i.e., urbanicity) when appropriate to contextualize teachers' interpretations of their self-efficacy beliefs.

Method

Sample

A final sample of 23 secondary teachers (4 of which were student teachers) participated in this study. When applicable, I distinguish between student teachers and beginning teachers, those in the first 1-3 years of teaching.

Middle and high school teachers were recruited via email using a stratified purposive sampling procedure which facilitates group comparisons (Miles & Huberman, 1994). Teachers were chosen primarily based on length of time teaching, but also with regard to content area, region of the United States, and urbanicity (urban, suburban or rural) to diversify the sample and thus include varying perspectives in the analyses. This resulted in an initial pool of 11 teachers. The snowball technique was then used (Miles & Huberman, 1994) according to which initial contacts are asked to provide additional contacts. Study participants were asked to recommend a teacher from the same school or teacher preparation program, ideally one that taught a different subject and had different amounts of experience. Four teachers were unable to recommend a matched participant in their teaching context and further recruitment efforts were made.

Teacher-reported demographic information (including teacher preparation) is shown in Table 14. Pseudonyms that begin with the same letter (e.g., Samantha and Stacy) represent teachers from the same school or teacher preparation program. I use the term "novice" teacher to apply broadly to teachers in the first three years of their teaching career.

[insert Table 14 about here]

The sample included a wide range of teaching experience, with more novice teachers (61%) than experienced teachers. Similar to national statistics in the U.S. (U.S. DOE,

2008), the majority of teachers were female (70%) and White (70%). Of note, only one teacher taught in a rural context

Procedure

Interviews were conducted in person in the teacher's classroom or university (n=14), on the phone or online via conferencing software (n=9). Three randomly chosen versions of the survey that differed only in the order in which items were presented were used to reduce order effects. Teachers' responses were audio-recorded and transcribed. After completing the 40-minute interview, teachers provided background and demographic information in an online survey. Respondents were compensated $40 for their participation.

Materials

During the think-aloud procedure, teachers responded to the short 12-item version of the TSES (Appendix A). The scale consists of three domains of teacher efficacy assessed by four items in each subscale: instructional strategies, classroom management, and student motivation and engagement. The response scale ranged from 1 (Not at all) to 9 (A great deal).

The demographic survey assessed teaching experience, type of teacher preparation program, amount of student teaching, perceived proficiency levels of their students (i.e., what percent below and at grade level), and the socio-economic status of their students (i.e., percent of students receiving free or reduced lunch).

Verbal Report Procedures

During the think-aloud portion of the interview, teachers verbalized their thoughts after reading the survey item (Ericsson & Simon, 1993). Cognitive interviewing techniques (Karabenick et al., 2007) were then used to target item interpretation and meaning. Teachers were asked to explain how they interpreted specific items as well as to elaborate on why they chose their responses to the items (e.g., the extent of agreement). Ad-hoc questions were asked for clarification. Appendix B provides the interview protocol.

Coding Scheme

All statements of a teacher about a particular TSES item were considered one unit of analysis for the coding procedure, resulting in a total of 276 units (hereafter referred to as teacher responses); 92 responses per subscale of the TSES. This approach has the benefit of not privileging the length of the response. One-third of the data were coded by three coders, including the author. When acceptable reliability was reached, the remaining two-thirds of the data were coded independently by the first author and the third coder. Interrater reliability for the entire data set was very good, with Fleiss' kappas ranging between .72 and .88 per code. Disagreements were then resolved by discussion until there was complete agreement.

The codebook was developed using an iterative process based on both a grounded theory approach (Corbin & Strauss, 2008) and the stated research questions. Warrants and evidence that teachers spontaneously provided for their efficacy rating were analyzed. For example, warrants for efficacy related to students being happy or angry after performing a teaching task would indicate that teachers based their sense of efficacy on effect of strategies on student attitudes.

Table 15 presents selected codes that were used to categorize teacher responses and develop broader themes: explicit reference to student outcomes, confusion about item meaning, listing of strategies to accomplish a task, and lack of controllability. Teachers' thinking and judgments underlying the assigned codes across items and within teachers were examined as well. Using grounded theory and the constant comparison method, I developed themes for the meanings of TSE, collapsing categories into broader themes when applicable. Teachers' responses to the demographic survey and responses throughout the interview were used to analyze the effects of the school context and teacher preparation on teacher thinking.

[insert Table 15 about here]

Results

First I present the extent that teachers based their sense of efficacy on evidence of effectiveness with students (student outcomes) across all domains, by domain, and by item within each domain. I then present findings in two sections: When do teachers consider their effectiveness with students and what are the different meanings and implications of their self-efficacy beliefs? That is, self-efficacy to achieve student outcomes. Second, when teachers do not consider student outcomes, what are the meanings and implications of their self-efficacy beliefs? This is self-efficacy to enact strategies. These two sections seek to determine whether this distinction matters for the meaning of teachers' sense of efficacy, in particular their use of research-based strategies known to be effective for student learning. As differences by urbanicity or length of teaching experience emerge, they are highlighted within each section.

General patterns of consideration of effectiveness. Overall, slightly more than half of the teachers' responses (55%) across all three domains of teaching tasks considered in the TSES (classroom management, student engagement, and instructional practices), included references to student outcomes as indicators of teaching success. These ratings of competence and success included, for instance, descriptions of situations when students "got it" or failed to understand content, when students enjoyed class, followed rules or when a strategy worked or failed to work with students. As shown in Figure 8, student outcomes were referenced as criteria for their efficacy judgments more frequently when responding to items from the classroom management efficacy (CME) and student engagement efficacy (SME) subscales (compared to the instructional strategies efficacy (ISE) subscale. This was expected since CME and SME items were more often constructed as student outcome-based efficacy items and ISE as teacher strategy-based items).

[Insert Figure 8 about here]

Within each domain of efficacy, teachers were consistent in their inclusion or lack of exclusion of student outcomes for individual items. One exception for the instructional strategies efficacy (ISE) scale was with the item *To what extent can you provide an alternative explanation or example when students are confused? (TSE for Alternative Explanation)*. In fact, relative to other ISE items, responses to this item were associated with more frequent inclusion of student outcomes (56%) shown in Figure 9.

[Figure 9 about here]

This was not the only ISE item that referenced students; *TSE to Craft Questions* also uses the word "students." Comparing *TSE for Alternative Explanation* to other items revealed that unlike other item, the task in this item presents a specific problem for which the

instructional strategy is needed: student confusion. This provides not only a specific purpose for the strategy being used, but a problem related to student thinking about the content that teachers need to solve. I refer to this hereafter as a problem-based teaching efficacy item, with the problem being located as something the students are doing or feeling. Importantly, teachers were not explicitly prompted to consider whether they could achieve the outcome; that is, how successfully they believed they could resolve student confusion. Despite this stipulation, teachers often spontaneously reflected that their alternate explanations and examples were successful in resolving student confusion:

> I haven't had a student ask a question … that I haven't been able to come up with a different explanation or example that the students are eventually satisfied with.—Katy, 5th year, mathematics

> We were just reading The Catcher in the Rye, and students had a question about Holden's feelings, and I felt like I was able to provide them an example that they understood.—Samantha, student teacher, English

Importantly, teachers had to feel responsible for influencing that student outcome; it was not solely achievement of the student outcome that positively impacted teachers' sense of efficacy. Hannah, a first-year teacher demonstrates this distinction:

> I find myself repeating myself and then students get upset. And so if a student is confused, I generally ask another student to explain it. I'd go with a "6" on that one, because while I'm not able to provide an alternate explanation all that often, everyone's almost always understanding what we're doing because of student explanations. And so I might go higher because we almost always understand, but because it's not me doing the explaining, I'll stick with a "6" [out of 9].

Hannah recognizes that the student outcome is achieved, but not because of her pedagogical skill to provide clear explanations. Thus her efficacy judgment is based on her skill to productively influence student thinking through alternate explanation, not just

achievement of the student goal. It should be noted that teachers, even student teachers, felt efficacious for this task, and ratings were generally high between 7 and 9 points.

In general, therefore, teachers considered student outcome when they were explicitly prompted to consider them, such as in the CME and SME items. It also appeared that teachers considered their effectiveness for one problem-based teaching efficacy item. For the remaining items, mostly within the ISE subscale, teachers did not explicitly consider their effectiveness with students when judging their sense of efficacy. The following sections consider the multiple meanings and implications of these two types of efficacy evaluations: efficacy for student outcomes and efficacy for strategies. In particular, the implications in regards to how the efficacy judgments intersect with teacher knowledge and implementation of effective teaching strategies is considered. As a note, efficacy ratings in this sample were generally high between teachers. Therefore, terms related to the strength of efficacy judgments (e.g., "high" or "low") are used as within-teacher rather than between-teacher comparisons.

Meaning of self-efficacy to achieve student outcomes

Responses that reflected teachers thinking about their self-efficacy to achieve student outcomes were grouped into four major themes: Controllability of the outcome, effectiveness in achieving positive outcomes with students, effectiveness of strategies in influencing student outcomes, and ability to solve problems using particular strategies (problem-based teaching efficacy). These themes are summarized in Table 126, which also includes patterns by years of experience and urbanicity.

[Insert Table 16 about here]

Controllability only. Some teachers' efficacy judgments reflected thinking of the controllability of the task in reference to what extent successful completion of the task would depend on factors within the classroom and under direct teacher influence. In other words, teachers considered whether the sphere of influence for achieving this outcome was situated within their classroom. When factors outside of the classroom contributed to success, teachers lowered their score by a point or two. In these cases, teachers either excluded or included evidence of their current effectiveness from their self-efficacy judgments, resulting in higher or lower efficacy, respectively.

Teachers who excluded evidence of effectiveness and rated their sense of efficacy highly were more often less experienced teachers. Naomi, a second year middle school teacher described her rationale for choosing an 8 out of 9 for the item, *How much can you do to control disruptive behavior?*

> I'd probably give this an 8 as something that I feel like I could—I can do a lot to control disruptive behavior. Am I doing that all the time, and is that always effective for me? Not necessarily, but I think that there's a lot that I can do.

Her effectiveness is separate from how much she feels she could do to influence the outcome of disruptive behavior. This allows her to rate her efficacy high as an 8 out of 9. Similarly, Brianna, a second-year English teacher, based her efficacy judgment on how much she was willing to try and learn, even though she currently struggled with student behavior. She explained that "every other week I'm trying to implement new strategies to control disruptive behavior" and that this task is something she knows "that I can get better at." For these teachers, the belief that they could learn more in order to positively influence that outcome meant that they rated their efficacy highly. Specifically, they do not talk about their skill in executing the strategies that are likely to achieve such outcomes. Naomi

expanded on her idea of influence versus control when she thought of her efficacy for motivating students, focusing on the word "help":

> I think that there's a great deal that I can do to help them and to support them in [valuing learning]. But to actually be sure that they are valuing learning or to know that's something that they've internalized for themselves—I don't think is as possible. But because it says "help them value," I would give that an "8", 'cause I think that there's a lot that we can do to foster that in students but not necessarily to control that [outcome].

Naomi careful articulate a difference in the word "control" and talks about the teachers' role to "foster" motivation in students. The phrasing was important for Naomi in this item, as she focused specifically on the word help. Similarly, Stacy, a student English teacher stated, "How much I can do [to control disruptive behavior] would be a great deal. How effective it is, which is a different question, might vary on the day, or the lesson… an 8 [out of 9]." Stacy may be evaluating a "great deal" based on beliefs of how much influence she believes she can have on the outcome, the variety of tools she has available, or her willingness to try and persist; regardless, like Naomi she rates her effectiveness in particular tasks as a separate issue from her teaching self-efficacy. *Their self-efficacy judgments therefore, likely reflect a judgment of how much responsibility for success is within teacher control (the upper bound of how much they can do as teachers) rather than a belief about whether they can execute specific strategies with skill in order to achieve the outcome*.

When teachers felt that success in tasks depended on the support of factors external to the classroom, their efficacy dropped. This most often occurred with student motivation items. Importantly, this could reflect adaptive or maladaptive beliefs about the possibility for students to be motivated. Naomi, the second year middle school teacher who felt

efficacious when she distinguished between controlling students' value for learning and influencing it, had a lower sense of efficacy "to motivate students who show low interest in school work." For this item, Naomi did not distinguish between controlling, helping and fostering outcomes:

> I believe that you can do a lot to help motivate students, but that it requires engaging a lot of different people. Involving families, engaging their friends, and a lot of different teachers and resources. And I think we have a lot of those structures in place at [School Name]. But I don't think it's something that I can control as much as something like disruptive behavior in my classroom. I think I'd probably go [with a] 6.

In response to this task, Naomi switches her thinking about the task as something that is dependent on the support of outside factors. Naomi lists many strategies that teachers can use to motivate students through drawing upon the resources and knowledge available in the community, an approach that is known to be particularly effective in urban schools (Gonzalez et al., 1995). Although she perceives that many of these structures to support her ability to motivate students are in place at her school and her feeling that teachers "can do a lot" for this task, she talks about controlling this outcome rather than fostering it. It's unclear why this item made Naomi think in this way; other teachers felt that valuing learning was more depending on outside forces rather than the immediate influence of the teacher. Her self-efficacy judgment is not based on how effectively she can involve families, engage friends and other teachers or what she can do to help students value learning (i.e., enact the strategies) but rather on her perception of how much control she has compared to tasks that take place within her classroom. While self-efficacy is by definition the perception of control (Bandura, 1997), Naomi brings up an important point. For urban teachers, the assumption that motivation largely depends on teacher actions

independent of their engagement with the school or outside environment may not be accurate. Capability to effectively engage with resources outside of the immediate classroom may be a crucial part of motivating students in urban contexts, thus changing the nature of the task of motivating students (Gonzalez et al., 1995). In fact, in his unpublished scale of TSE, Bandura included efficacy for engagement of community resources, but these items were not included in the development of the TSES for lack of relevancy (Tschannen-Moran & Hoy, 2001). Although this thinking about external resources was most salient among teachers in urban contexts, other teacher identified this as a difference between elementary and secondary schools or the fact that students are older and more independent in middle and high school.

While Naomi's response indicates a belief that students can be motivated in her context because of resources in the environment, other teachers in urban contexts exhibit maladaptive beliefs about the influence of home environments on student motivation and behavior. These teachers' attributions for their failures to reach students are representative of what has been called low general teaching efficacy, or belief that teachers can do little to overcome negative environmental influences (see Chapter 2 for an extensive discussion). One such teacher was Florence who was in her ninth year of teaching at the middle school level in an urban setting. She revealed that she has continually struggled with classroom management because being "hard" is not "authentic with her personality." In response to her struggles, Florence mentioned that her principal re-assigned her from general education to gifted and talented students, assuming these children would exhibit less problematic behavior. Even still, she continued to struggle with management and motivation, and

throughout the interview made repeated references that some students could not be reached:

> *I can* do a lot but there are some kids that, who, whose life circumstances are such that, almost no matter what I do, I'm not going to be able to motivate them. Because they've got such stresses outside, or other pressures are coming to bear on them, or ability issues. It's a difficulty that I'm unable to overcome as a teacher.

Like Naomi, Florence's efficacy judgments are based on her sense of control in the situation. Unlike Naomi, Florence response focuses only upon her influence on student motivation, locating the task of motivating students entirely within her classroom, and thus upon her shoulders. She does not see strategies by which students might successfully be motivated; to her it is an impossibility because of detrimental external pressures. Florence does not think of resources in the environment that might help her to motivate students. Given that her school principal's chose to give her "easier" kids rather than support her to develop her management skills, it may be that those supports do not exist in her environment. Lack of support may have contributed to this belief that she will not be able to reach some kids as well as viewing the task as resting squarely upon her shoulders. Florence's school context is contrasted to Naomi's school in which she states that she always has support to help solve problems related to students.

> If I go and I sit with Rebecca, my planning partner, or Maria, my family partner ... so much of the conversation at [school name] is about strategies . . . You do some deconstructing of the problem, but it's so focused on your solution, and what are your next steps, and where are you going, that I feel like I don't necessarily do as much reflection on my confidence . . . It's more of, just a question of, "What am I going to do about it now?"

As a second-year teacher, Naomi's school has provided her with two colleagues to support her to find solutions, a "planning" partner who also teaches 7th grade science and math and a "family" partner that teaches English and Social studies to her students. It is clear that

she meets with these individuals often and that they help her solve specific problems with students rather than remove the problem students from her classroom, as was done in Florence's case. Naomi's thinking about finding solutions to problems and confidence not being a consideration are remarkably similar to Dennis, a suburban high school teacher with 25 years of experience. Dennis explained:

> It's not a matter of doubting my capabilities. It's a matter of 'Have I found all of the different ways . . . in order to be successful?'

Either with support in the environment or through the knowledge gained from experience, teachers in this sample who focused on problem solving were also less likely to even consider their confidence about whether they could influence student outcomes.

Importantly, both Naomi and Florence rate their sense of efficacy for motiving students lower than for other tasks, as a 6 out of 9. Their reasons for their low efficacy, however, translate to entirely different beliefs about the possibility of the outcome and who might be important in supporting, rather than taking away from, student success. Naomi's response indicates that low efficacy for the outcome of motivating students may mask an important belief that teachers, in concert with the community resources, can have a positive influence on student motivation. Identifying and engaging resources in communities is a research-based strategy to support student performance and motivation, particularly in under-resourced communities that are common in urban locations (Ladson-Billings, 1995; Weiner, 2006). Teachers with these beliefs, like Naomi, are likely to exhibit vastly different behaviors and attitudes towards students both in and out of the classroom. For TSE for student motivation to have practical utility, TSE measures should

be able differentiate between these types of high and low efficacy for motivating students, particularly in urban contexts.

Effectiveness with all students: Experienced teachers. Finally, efficacy as controllability of the task for more experienced teachers, however, was more often related to a lower sense of efficacy and greater focus on effectiveness with students. These teachers usually reported experiencing success with students yet rated their efficacy lower when they reflected on rare instances in which they failed to reach specific students. Katy, a fifth-year mathematics teacher in an urban setting reflected at length about one student "intent on being disruptive" that she could not calm, and she rated her efficacy for this task as a 6 out of 9; however, she acknowledged that she was only thinking of about "one [student per] year so far." Such thinking was not limited to urban teachers; Dennis is a mathematics teacher in his 26th year of teaching in a suburban, affluent high school. Dennis first lists a number of strategies he uses to motivate students yet rated himself a 7 out of 9, adding, "[I'm] being hard on myself at that one." When asked to clarify this statement in follow up questioning, Dennis replied,

> Because there's always one or two students that you're like, 'Thaagh! I didn't get them.' So, I would say I'm probably hard on myself there, because I want to get those two, and I can't get them.

Although Katy and Dennis consider themselves largely successful in motivating students and controlling behavior their lower scores reflect a focus on "exceptions to the rule," or the somewhat isolated cases in which these teachers could not reach a particular student. Dennis states, "If there's one or two students that you didn't get, you feel like you're unsuccessful. But that's kind of what makes you successful." Importantly, these experienced teachers' high standards are coupled with an awareness of the inability to be

effective with every student. This awareness lowers their efficacy score by a point or two, while for Naomi and Stacy lack of effectiveness was a "different question" from their sense of efficacy for the task. The perception of failure for Dennis and Katy is tied to nearly unattainable high standards, however: feeling unsuccessful because of on relatively rare cases in which they were not successful with particular students. *Although Katy and Dennis consider themselves largely successful in motivating students and controlling behavior their lower self-efficacy scores (although only by 1-2 points) can be interpreted as an acknowledgment of the uncertainty of the outcome (e.g., 100% success) rather than a belief in their personal skills to influence the outcome among most students.* In essence their "desired level of competence" (Tschannen-Moran & Hoy, 2001, p. 210) is to reach every student yet the fact they will not be able to reach all students is also salient to them. This may indicate a reference bias (Duckworth & Yeager, 2015) in which more knowledgeable and experienced teachers have higher standards for success impacting their self-ratings. Dennis again demonstrated this reference bias in response to an item that asked about his self-efficacy to help students believe they can do well in "school work." Dennis rated himself a "5" because he interpreted school work to mean in other classes besides his own. "if I was going to be a phenomenal teacher, I would be able to not only make them be successful in math, that would also inspire them to be successful in other areas." No other teachers thought the item referred to work outside of their classes.

As seen in this section, not all teachers spontaneously thought of specific strategies when asked to think about their capabilities to achieve student outcomes. The following section considers the meaning of efficacy beliefs when teachers did think of specific strategies, either spontaneously or when prompted to by the item, and their effectiveness in

achieving desired student outcomes. Particular attention is paid to whether the strategies teachers generated were reflective of research-based strategies known to be effective for classroom learning.

Self-efficacy as effectiveness of strategies in influencing student behavior

Another theme that emerged when teachers thought of achieving student outcomes was consideration of strategies, or the principles that underlie them, to achieve the student outcomes indicated. Analysis of responses revealed potential misunderstandings or lack of knowledge of the best strategies by which to achieve certain outcomes, particularly in relation to controlling disruptive behavior. The differences in these were somewhat related to teachers' years of experience which may have functioned as a proxy for level of pedagogical knowledge.

Three experienced teachers and one student teacher spontaneously articulated broader principles of classroom management that work to shape the learning environment in response to *TSE to Control Disruptive Behavior*.. I highlight two teachers, Amanda and Dennis, with 8 and 25 years of experience, respectively. Amanda, a high school English teacher, talked about the teachers' role in setting expectations:

> I'm going to jump to an '8' [out of 9 on the response scale] here. I think the teacher has a huge hand in shaping the classroom culture and expectations … for what is okay to do in the classroom and expectations for what is not okay to do. Very much so I think teachers set that tone from the beginning of the school year.

Amanda talks about the importance of the beginning of the year and setting expectations. While she does not go into detail about how she accomplishes establishing a positive "classroom culture" and setting the right "tone" from the beginning of the year, this may be

because she does not need to extensively analyze the task because of her experience with it (Tschannen-Moran et al., 1998). Similarly, the most experienced teacher in this sample, Dennis, mentioned setting the right "tone" in response to *TSE to Control Behavior*, not elaborating on how he accomplished this or indicating specific instances of success or failure. Moreover, Dennis mentioned the classroom environment again when responding to the other CME item, *TSE to Calm Student*:

> The number one thing that you do to calm a student, or to calm a student who's disruptive or noisy is to incorporate things so that that does not happen in the first place. You create a climate so that that will not happen in the first place. Student teachers ask me, "Well what do you do when somebody is standing on their desk and they're sitting there throwing papers at somebody else?" And I guess my question is, "What did you do to prevent that in the first place?"

Dennis connects the two tasks with an underlying principle of creating and maintaining a strong and productive classroom culture. For these experienced teachers, this principle is most salient when they respond to the item, rather than a list of discrete strategies by which to achieve the outcome. This may indicate that because of their knowledge and experience, these teachers are more able to extrapolate the underlying theories of the strategies (Korthagen, Loughran, & Russell, 2006) which contributes to a higher sense of efficacy. As Dennis points out, student teachers that he mentors are unlikely to think about creating a positive classroom culture as a way to control disruptive behavior; so were most teachers in this sample, regardless of their years of experience.

Unlike Dennis and Amanda, the majority of teachers with less experience thought of strategies to respond to misbehavior after it occurs rather than prevent it from happening, and did not always think of strategies that are known to be effective. After reading the item, Ally, a middle school teacher in her fifth year, immediately rates her self-

efficacy as an 8 out of 9, saying "I can do a lot." As she continued to think aloud, some strategies slowly come to mind:

> I can control disruptive behavior because I can … I don't know. I can send them out, I guess. [That] is controlling, getting it out of the classroom. Or calling home works pretty well. And also I do a lot of positive [reinforcement].

Two of the three strategies that Ally spontaneously generates are punitive and reactionary in nature: sending students out of the classroom or calling home. Neither of these strategies indicate thinking about the classroom environment as a whole. Importantly, Ally recognizes her lack of effectiveness and continues to think, "I would change that to a '7' because I don't think I'm that great. I think I've improved a lot as a classroom manager but I think that's a piece I still need to work on." Accordingly, she lowers her score by one point to a 7 out of 9. Rather than think about the controllability of the outcome, Ally thinks about her overall effectiveness as a classroom manager. Given that the strategies she eventually offered are not known to be particularly effective for either controlling behavior or creating a structured classroom environment, the lowering of her efficacy scores indicates she is aware that she could be more effective.

Teachers in their first or second year were more likely to rate their sense of efficacy for this task even lower than Ally, even though they referenced the same strategies. Justin, a first-year science teacher, also responded that he removes disruptive students from the classroom or re-directs them, but rated his sense of efficacy as a 5 out of 9. Likewise, Hannah, a first-year teacher, rated herself as a 4 and explained her difficulty in managing disruptive behavior when it occurs at the classroom level, "It's really hard for me to get everyone back together if everyone's off task." She clarifies that if the question were to ask about controlling behavior with just "a few students … then [my self-efficacy rating]

would be much higher [than a 4 out of 9].” Importantly, Hannah identifies that this task is asking about the classroom as a whole rather than individual students. She attempts to employ the same strategies for each task, yet recognizes they are not as effective on a large scale. Similarly, Justin, explicitly related *TSE to Control Behavior* to the other disruptive behavior item, *TSE to Calm Student*: “I feel like that’s the same thing as calm a disruptive or noisy student.” For these novices, the connection between the items is in the use of similar strategies that are reacting to, rather than preventing disruptive behavior. For Dennis, the connection between the two tasks was preventing disruptive behavior through controlling the parameters of the learning environment. Thus, novices were less likely to spontaneously identify either the principles that work on a larger level nor the discrete strategies that would be effective in managing the learning environment. Their lower sense of efficacy, therefore, seems to be a calibrated reflection of the effectiveness of their strategies and lack of knowledge about how to do this task more effectively.

Some beginning teachers did think of strategies that are likely to be more effective in influencing the student outcomes, but their lack of efficacy stemmed from the inability to execute strategies at a desired level of proficiency. For example, Crystal mentioned “classroom management … it’s just the ability to execute a lesson that’s hardest.” When asked to clarify, she stated,

> How you execute a lesson. How much time to give, pacing. A lot of times I go in and observe other teachers or I reflect on my own, but that’s not something I’ve been taught … in my master’s program … I learned how to lesson plan, basically.

Crystal differentiates between being able to plan a lesson and being able to execute a lesson in the classroom, differences between planning for teaching and the interactive components of teaching (Jackson, 1990). Crystal identifies this as a component of

classroom management, showing knowledge that managing the classroom also depends on managing the pacing of content as well as behavior (Klieme, 2012). Despite this greater knowledge of what is needed to manage a classroom, Crystal still struggled with pacing her lesson, adjusting her management plans for each class. Thus Crystal's sense of inefficacy stems from inability to execute the strategies that she knows are important for student learning.

This distinction whether teachers' low efficacy stemmed from lack of knowledge of effective strategies or lack of skill in executing strategies is important because such differences may have implications for the type of professional support that would be most effective in both improving teachers' sense of efficacy and the quality of their instruction. With her knowledge of the interconnectedness of content and management, Crystal would benefit from support based on observation of her pacing of lessons. Ally, Justin and Hannah, teachers who did not spontaneously see the connection between content, the environment and student behavior would likely benefit from observation focused on helping them see how support in seeing how structures, routines, and content contribute to management of student behavior. Although Ally is more experienced than the other teachers mentioned, her thinking is more similar to the less experienced teachers. This provides support for the earlier assertion that factors such as pedagogical content knowledge or skill with self-reflection (which may or may not increase with time in the classroom) may be driving differences among teacher thinking rather than years of experience. On the other hand, Ally's response may also indicate that the longer teachers spend in the classroom, the less likely they are to rate their sense of efficacy lowly because they have seen improvement over time. Although Ally acknowledged she was "not great"

as a classroom manager, she focused on her improvement over time and rated her efficacy as a 7 out of 9. Thus, more time in the classroom may contribute to higher efficacy because of improvement, though teachers may not be reaching external standards for quality teaching.

Ability to solve problems using particular strategies. Finally, teachers also thought about whether they achieved student outcomes when they thought about their ability to execute specific strategies in response to a particular problem exhibited by students. These types of responses were most common for efficacy items, which specified a student problem that needed to be solved, or problem-based teaching efficacy items.

As described previously, there was one ISE item, *TSE for Alternative Explanation*, more frequently associated with consideration of effectiveness with students. This item asks teachers about their confidence to enact a specific instructional strategy—providing an alternate explanation—in response to a student problem—confusion. The strategy is specifically related to a research-based component of effective teaching: clarity of expression and understanding of content. In this item, teachers did not have to generate knowledge of a research-based strategy to achieve an outcome like other items asked them to do (e.g., *TSE to Control Disruptive Behavior*). As a result, high self-efficacy for *TSE for Alternative Explanation* relative to other ISE items was more often a reflection of higher confidence to successfully enact a strategy in response to a student problem, with success defined as achieving the student outcome.

One SME item, *How much can you do to calm a student who is disruptive or noisy?* did show the same pattern of meaning as *TSE for Alternative Explanation*: efficacy ratings more often meant high efficacy to reach students through their own skills. Similar to *TSE for Alternate Explanation*, teachers were asked to consider how to respond to a particular problem related to students: being disruptive or noisy. Unlike *TSE for Alternate Explanation*, teachers were not prompted to think of any particular research-based strategy to respond to the problem.

Despite this, teachers more often reflected on the strength of the relationships that they had with students and their ability to diagnose why misbehavior was occurring as important factors in being able to calm students.

> I somewhat agree, depending on your relationship with that student. Unless you're able to kind of like communicate, or just get through to that student, which not always the case … generally speaking I'm not as good at talking down [but] there's this one student here who—I think I'm like the first teacher to get along with her. I'm able to talk her down from being just really upset.

Caleb reflects on his relationships with students, acknowledging that "talking" or calming down isn't one of his strengths. He acknowledges that he can successfully talk down a student with whom he has a relationship. Although having a personal relationship with students is not mentioned in the item, teachers—particularly more experienced ones—seemed to implicitly recognize that this is necessary, such as Patricia who states that "understanding why they're behaving in that way is important … as the school year progresses you can figure out how to deal with [that] student appropriately" and Dennis who states that he "can recognize their personalities." Getting to know students is important, as well as diagnosing why misbehavior is happening. It seems that the problem presented—a disruptive student who needs calming—prompted teachers to consider the

relevant skills and knowledge needed *in order to* solve that problem successfully. Strong teacher-student relationships are an important characteristic of effective classrooms, yet this is typically categorized under Supportive Classroom Climate rather than Classroom Management (see Chapter 2 for an extended discussion).

Further demonstrating that presentation of a problem had an impact on the meaning of teachers' self-efficacy judgments, teachers who had, in other items, evaluated their SME or CME efficacy as controllability of the outcome without reference to specific strategies, did not do that for *TSE to Calm Student*. Recall Florence and Naomi, the two teachers who thought that motivating students to do well was a factor out of their control, albeit for highly different reasons. Florence acknowledged she was not confident in classroom management and also reflected beliefs that she could not control the behavior of certain students:

> I don't feel very confident about it… but I can do more to control disruptive behavior than I actually do [and] the question is how much "can" I do. … I'm going to go with a 7 because … I could do quite a bit more than I do. It's lower than a great deal because there's some kids that I really can't control regardless of what I do personally or even what the office does.

Florence admits that she is not confident here and that some students cannot be controlled—though she gives herself a high (relative to other scores) rating for this item. Similarly, Naomi also recognized that she was not effective but based her sense of efficacy on controllability, giving herself an 8 out of 9. Florence and Naomi consistently rated their sense of efficacy for classroom management tasks in this manner—focusing on what they could do—yet showed a different response to *TSE to Calm Student*. In response to this item, Florence based her self-efficacy judgment on how strong her relationship is with the student like other teachers:

> Depends on the kid … That's such a relationship question in terms my pull with the kid, depending on how strong our relationship is, how well I know the kid, and how much he or she respects me … that's clearly one where it depends. A 5.

Florence does not reveal the same maladaptive belief that some students just could not be controlled; instead she judges how well she has established a relationship with the student she is attempting to calm. Recognizing that she does not have a strong relationship will all students, Thus, her judgment for this task is based not on how much she "could do" if she wanted, nor does she indicate that this task is beyond the realm of her influence as it was in other classroom management items. Her rating reflects actual confidence that she has the skills to positive influence (i.e., calm) students through relying on her relationship and knowledge of that student. Similarly Naomi felt less confident and effective with this task of calming students:

> I feel like that's one that I'm not always that effective at because it's hard for me sometimes to gauge what student needs in terms of how to calm them. And what's really at the root of their disruption and what strategies to use. There's a lot that I can do, but … it's not something that I'm always doing 100% effectively.

Naomi refers to the knowledge that other, more experienced, teachers identified—the ability to diagnose why a student is misbehaving as important in solving the problem. Crystal, the teacher with low confidence for classroom management, also referenced the importance of figuring out why students were misbehaving, indicating that this item "is a little bit different." Thus, relative to other items, teachers were more often thinking about two necessary skills that they could develop: relationship building with students and diagnosing reasons for a student's misbehavior. In all, this reflects greater responsiveness to students, which are typically characteristics of creating a supportive classroom climate. Moreover, it appears that this item did not inspire teachers to think of controllability, or

rather teachers implicitly assumed that this task was well within the realm of teacher influence, allowing them to evaluate their skills in solving this problem of student misbehavior.

Interim Summary of Efficacy for Student Outcomes

This section explored the meaning of teachers' self-efficacy judgments when they consciously considered student outcomes as the task. Two patterns were identified: teachers thinking of the controllability of the outcome and teachers thinking about what strategies they could use and how effective they were at achieving the outcome.

Differences within each pattern were found by years of experience, urbanicity, and also by item. Experienced teachers were more likely to rate themselves lower because they had higher standards for success with all students relative to other teachers; high standards were tempered with the acknowledgement that it was a virtual impossibility to reach all students every year. Less experienced teachers did not acknowledge that impossibility, instead evaluating their sense of efficacy based on evidence of improvement over time or how much the task was centered within their classroom and under their complete control. Urban teachers were more likely to reflect on the environment as a negative influence on student motivation and behavior, though one teacher recognized that resources in the environment existed that could support student motivation; despite this, her efficacy rating was low because those supports were external to the classroom.

Finally, teachers did not always reflect on research-based strategies to achieve student outcomes, particularly less experienced teachers, and some identified problems with execution of the strategy rather than knowledge of the strategy. One item more

successfully elicited teacher thinking about research-based components of creating a supportive classroom climate, thought this item was meant to a be a part of the classroom management scale. Such variation in the meaning of efficacy beliefs although teachers were considering the task in similar ways casts doubt on our ability to make practical and logical recommendations for ways to improve teacher sense of efficacy. Building on these findings, the following section explores the meaning of teacher efficacy when teachers did not consider the effectiveness of their teaching strategies on student outcomes.

Meaning of self-efficacy to enact strategies

When teachers were not evaluating their efficacy to achieve student outcomes either through reflecting on the controllability of the outcome or the effectiveness of particular strategies in reaching students or solving problems, their self-efficacy beliefs were more often based on a) the frequency with which they enacted strategies and b) the range of options or knowledge of strategies that they possessed or c) how well their use of strategies aligned with personal standards for quality. These types of responses occurred most often in response to instructional self-efficacy (ISE) items. It is important to note that teachers more often expressed being confused over item meaning, particularly in relation to the terms "alternative strategies" and "good questions." All but one teacher resolved their confusion regarding the meaning of through thinking aloud. Appendices A and B summarize interpretations of each task; relevant interpretations as to the changing meaning of TSE are highlighted here. Table 17 displays the multiple meanings of TSE when teachers did not think of student outcomes, largely in response to items from the instructional self-efficacy scale.

[Insert Table 17 about here]

Self-efficacy as variety of strategies known and/or used. Teachers were inclined to judge their sense of instructional efficacy by thinking of how capable and how frequently they implemented a variety of different strategies in their classrooms. This was an expected response to item that specifically asked about variety, *TSE for Variety Assessments*, ("How much can you use a variety of assessment strategies?"), but was also seen as the meaning of teachers' self-efficacy for *TSE for Alternative Strategies* ("How well can you implement alternative strategies?"). Marcus, a third year teacher was confused about the stem "how much" in the first item and wanted to know what time range he should be thinking about as he considered the variety of assessments:

> I'm trying to process this one in terms of the different types of assessment strategies. I don't know, "how much," if it's asking me how often, or if it's what is the range there? I'd also say that for this question it'd be helpful if I knew, within a week, or within a month or something.

In follow-up questioning, Marcus clarified that the question made him think about the different ways that he had assessed students in the past week, rating himself highly because he uses a variety over a given week. His TSE, therefore, is based not just on how capable he is of implementing the assessment strategies but how varied his assessment strategies are over a given week. Dennis, the experienced mathematics teacher, was also confused about the item intent for using a variety of assessments:

> 'How much' can I *use* a variety? Once again, it's a quantitative question. Is it asking how many assessment strategies do I *know of*? Is that kind of what it's asking?

Dennis identified "how much" as a quantitative, perhaps influenced by his discipline of mathematics. Although the item asks about use of strategies, Dennis is initially confused whether this is the same as "knowledge" of strategies. It may be that the distinction between use of strategies and knowledge of strategies is useless to Dennis, particularly given that he had high self-efficacy for most tasks. Accordingly, throughout the think-aloud, Dennis tended to list off the different strategies or "tools" that he had to accomplish the tasks. Naomi referred to knowledge as the "range of options" she had in order to accomplish the tasks which was distinct from effectiveness:

> These [items] that said 'how much can you do.' When I hear that, it sounds like "what's the sort of quantity of options?" Like, how many different things can I do. Versus, "How well can you control disruptive behavior" tells me more about like how effective am I at doing something like that.

While Naomi identified the item stem "how much" (which was present in 8 of the 12 items) compared to "how well" as influencing her thinking towards a "quantity of options," it was only responses to items with words such as "variety" and "alternative" that most frequently reflected teacher thinking of variety, knowledge, range of options, such as Bob who said, "I have quite a variety of assessment strategies available to me." This pattern of thinking about this item was true regardless of whether teachers did consider the effectiveness of their strategies at other points during the think-aloud, suggesting it was the task or item phrasing that influenced teacher thinking.

Although *TSE for Alternative Strategies* did not specifically prompt teacher to think of the variety of strategies they used, teachers rated their sense of efficacy for this task based on similar criteria of evidence of variation. Interestingly, reasons for varying strategies were less frequently stated by student teachers, such as the three below:

> This is kind of take me back up to number five about the "variety of assessments," thinking a variety of teaching strategies in the classroom – Samantha
>
> I feel like that's probably one of my strengths is my ability to, to vary what happens in the classroom in terms of instruction –Stacy
>
> How well you vary your instruction ... [in terms of] lesson planning –Darren

These student teachers and a few beginning teachers were less likely to explicitly name why one might want to vary instructional strategies compared to teachers in the classroom who identified reasons related to different ability levels of students. By simply including a variety of strategies in each lesson, these teachers seemed to be confident that they were mastering the task presented This may indicate that the purpose of using alternative strategies was not salient to some novices and may need to be explicit in the item.

For in-service teachers, they more frequently related purposes for varying strategies, yet they did not always consider reasons based on scientifically sound ideas about how students learn. This was most evident in response to the item *TSE for Alternative Strategies*, in which 11 teachers (48%) expressed confusion about what constituted an "alternative strategy." The majority of teachers interpreted this to mean differentiated instruction, or ability to adjust instruction for learning needs of different subgroups of students in the classroom (n=11). Seven of those teachers thought of subgroups based on students' perceived learning style (n=7), a perceived student capability widely believed to be stable over time, although little empirical evidence supports this theory (Willingham, Hughes, & Dobolyi, 2015). Victoria, a first-year teacher, thought about "the visual learners, the auditory learners, and all those different kinds of learners." She stated that "you always read about" learning styles, suggesting that prevailing

discourse about the importance, and stability of, learning styles may have influenced her interpretation of this item. Hannah reflected that learning to use "different modalities" was a focus of her master's program and so, "We move around, we do gallery walks to try and get our kinesthetic learners in there." Hannah suggests that learning styles theories are currently being taught in her master's program, and Dennis provides evidence that even highly experienced teachers held this beliefs:

> I can show different methods and I can be … more verbal or visual. I can be more left-brain or then some days be right-brain. I can be more cognizant of what they're thinking so that I can adjust to how they, as a class, end up thinking and how they learn best. –Dennis

Dennis and other teachers felt efficacious when they could and did present content in multiple different ways[4]. This belief in learning styles was evident among novice and experienced teachers like Dennis, speaking both the prevalence of this belief among educators and the need for ongoing professional development for teachers focused on research-based practices.

Recall that in the previous section that teachers often generated strategies to achieve outcomes that were not aligned with research-based strategies (e.g., TSE to *Control Behavior*); however they seemed to realize that these strategies were ineffective with students and thus rated their sense of efficacy lower. For strategy-based efficacy judgments, however, teachers did not tend to rate their sense of efficacy lower or higher based on evidence of effectiveness of those strategies on student behavior or learning. Reasons given were related to improvement in use of different strategies over time, time or curriculum restraints, or knowledge of strategies gained from professional development.

[4] Teachers who didn't think of subgroups by learning styles (n=4) thought about other ways to differentiate instruction besides presenting content, such as assessment.

A few teachers who did, however, consider the impact of their strategies on students did not always think of the strategies' impact on student learning or cognition, as might be expected from an instructional strategy. Instead, they were more likely to consider student enjoyment, interest or behavior. These tended to be novice teachers, like Victoria and Vivian:

> I did a lot of this kind of stuff, and I think that that was effective. And the kids told me that they found that to be very good, and they'd never really seen much of that before. So I would say "8" – Victoria
>
> I need to still practice, especially in math, not falling into that same routine of giving notes, doing examples, and then giving them practice problems. Because it's really easy to do that, and it can get really boring that way – Vivian

Victoria seems to be thinking about student enjoyment based on the novelty of the idea and Vivian about student interest. This type of response compares to Amanda, a teacher in her 8th year who thought about trying a new strategy of student book clubs for the first time. After talking about her willingness to try new things, she states, "the [students] were able to do what we asked for them to do." Amanda doesn't talk about interest or enjoyment but whether or not she was able to help students complete the desired task. While it is somewhat unclear whether students learned what was intended from the activity, compared to the novices Amanda seems to focus more on what students were doing with or as a response to the strategies, rather than their enjoyment or interest in the strategies.

Also absent from all but one teacher's response was consideration of how appropriate the strategies they used were based on the instructional context or ability to use those assessments to inform further instruction, indicators of effective use of assessments (S. A. Cohen, Hyman, Ashcroft, & Loveless, 1989). Brianna, a first-year teacher, was the

only teacher who identified something near to these issues when she said, "I think it's more about knowing if I'm using a correct one for a particular assignment. I feel like I do use a variety of them but I just don't know if they're working." Brianna mentions of a particular assignment hints at the idea of content playing a role in deciding which assessment to use as well as an inability to interpret the information gained from the assessments ("if they're working). While Brianna rates her sense of efficacy low (4 out of 9) because she is unsure about the appropriateness or effectiveness of her assessments, it is unclear whether other teachers who rated their sense of efficacy high for this task were also confident that their assessments were appropriate. Thus teachers may be confident in their use of assessments, but not using them appropriately to leverage their understanding of student thinking and to determine further instruction.

Self-efficacy to meet standards for performance. Another common meaning of strategy-based self-efficacy judgments was efficacy to perform or develop content to meet standards for quality performance. These standards could be contextualized or decontextualized; in other words, dependent on the instructional context (including students) or not. This was most clearly seen in response to the item "To what extent can you craft good questions for your students?" (*TSE to Craft Questions*) likely because the word "good" requires that teachers implicitly or explicitly define what constitutes a good question against a particular standard. No other item contained an evaluative or subjective word to describe the task; the use of the evaluative word "well" is confined to the stems of items.

Most commonly, teachers considered a good question against a decontextualized standard: a "high level thinking questions" that requires deeper levels of understanding to

answer (n=12). A few teachers, many from the same alternative teacher preparation program, specifically mentioned Bloom's Taxonomy (Krathwohl, 2002). This static view of what constitutes a "good question" in any situation was associated with lower efficacy scores when teachers felt they could not use these questions with all students or in all situations. Justin, a first-year science teacher, when asked how he decides if a question is good, clarified, "I use Bloom's Taxonomy, thinking about those different levels." Justin's efficacy belief for this type of question is related to his ability to craft and use those questions in a variety of contexts:

> So I would probably say [my efficacy scores is] about a "7 [out of 9]", because while I feel pretty confident during lecture, I still have improvement to be made, especially for quizzes and tests.

Justin's sense of efficacy is attributed to his ability to create a question meant to inspire higher-level thinking rather than specifically how effective these questions were at inspiring higher-level thinking among students. More significantly, he is sure that he knows what a good question in every context—one that aligns to higher tiers of Bloom's Taxonomy. Justin's idea of a good question is de-contextualized in that he does not indicate that other types of questions might be considered good depending on the instructional context; for example, a "good question" as a lower-tier question that reveals a lack of knowledge that is prohibiting higher-level thinking. Similarly, Ally thought of good question as higher-level but had low self-efficacy for this task because she could not implement these questions with all students:

> My higher-level kids I think that I can do this with. But lower-level kids, I don't necessarily know that I'm getting to those higher-level questioning strategies.

Ally thinks of failure when she doesn't ask higher-level questions to her "lower-level" students. In other words, the meaning of low efficacy for higher-level questions for Justin and Ally was inability to design or implement these good questions in all situations or with all students.

In contrast, other teachers defined "good questions" as valid diagnostic or assessment tools, with emphasis on validity whether the question was successful in eliciting its intended purpose, i.e., evidence of student understanding, thinking, or discussion. Marcus, a science teacher in his 3rd year, defined good questions as one that "elicits student understanding of the specific content … it's a valid assessment tool." Vivian, a pre-service mathematics and social studies teacher, similarly stated that the purpose of a question is "to assess what your students know." These teachers' responses did not qualify what type of thinking (i.e., high-level) must be demonstrated by students in their responses since these questions were seen as diagnostic tools which could be used to determine *at which level* students were thinking. Teachers, therefore, did not feel inefficacious because they were unable to implement these questions with all students or all the time, but when the questions failed to achieve their purpose. For these teachers their self-efficacy for crafting "good" questions was dependent on student responses to those questions:

> I've struggled with … making sure that I always get the right response. I feel like a lot of times there are responses that are not anywhere near what I was looking for. It makes me rate this one a little bit lower than others.

Rather than feel he cannot craft good questions for all students or use them in all situations, Marcus focuses on whether the intent of the question is actualized as seen through students' responses. These findings suggest that thinking of questions as diagnostic

assessment tools was more closely associated with thinking of the instructional context and a contextualized definition of "good." These teachers thought about the context of what was being learned, how students were thinking about it, and evaluated their efficacy to gather that information from students through questioning: the interplay of those three things constituted a good question. Two teachers who thought of questions like this (Victoria and Vivian) came from the same teacher preparation program that highly emphasized developing good questioning strategies; from their responses, there were similarities in the way they thought of good questions as ways to elicit student thinking and knowledge.

In sum, two different meanings of self-efficacy for strategies emerged from the data: 1) confidence in the variety of strategies that they were capable of using –and did use—in their classroom, and 2) confidence in ability to meet standards for performance that were contextualized or decontextualized. More subjectively interpretable words such as "good" or "alternative" led to great variation in the types of strategies that teachers considered and thus, different patterns of meaning related to which types of strategies teachers felt efficacious to implement and why. Again, evidence emerged that teachers' efficacy beliefs were based on beliefs about teaching and learning not substantiated by scientific research: learning styles. However, because teachers did not spontaneously reflect on how students responded to strategies associated with these beliefs (typically using a variety of modalities), it was not clear that these instructional self-efficacy judgments were always calibrated to effectiveness with students in terms of learning. In fact, when teachers did spontaneously think of the effect of their instructional strategies on students, they often thought of student interest or enjoyment rather than cognition or

learning. This seemed most evident in novice teachers. Furthermore, the lack of instructional purpose in these items (e.g., purpose for using alternative strategies) was particularly problematic for student teachers who were less likely to spontaneously consider why they might be using these strategies. Finally, a small minority of teachers did consider the instructional purpose of strategies, but only when they interpreted "good" as a contextualized standard dependent on students.

Discussion

This study sought to reveal potential differences in the meaning of TSE beliefs based primarily on thinking about efficacy for student outcomes versus enacting teaching strategies. Overall, for many tasks of instruction, management and student engagement, teachers considered how students responded or would respond to their efforts as they judged their sense of efficacy. However, this was done in vastly different ways and with different implications for both the meaning of TSE beliefs and the importance of item construction.

One major finding was that teachers' self-efficacy often meant how controllable teachers believe the student outcome was; that is, whether primary responsibility for success with students resided within the space over which teachers' had the most control and responsibility: their classrooms. This meant that TSE for classroom management was high among teachers in this sample, consistent with other studies (Brouwers & Tomic, 2000; Ryan et al., 2015) but not because teachers always felt particularly effective at implementing particular strategies to achieve the outcome. This finding raises an important question: How practical and useful is an assessment of TSE that means "I could do a lot if

I were, or when I am, more skilled." Simply by analyzing the nature of the task of managing norms for behavior in classroom where one is a leader, it is plausible to predict that teachers feel they have more control over that versus affecting a students' internal motivation to learn. Clearly this is related to teacher burnout and willingness to stay in the profession, but it seems prudent to move beyond general outcomes of retention and happiness to what TSE means for teachers' instructional practice: how might TSE predict use of specific teaching behaviors known to be effective for student learning?

Building on this question, it was encouraging that the meaning of TSE *was not* consistent within teachers; teachers didn't always rate their TSE by how controllable the outcome was. For instructional self-efficacy and certain tasks that were phrased as problems to be solved in the classroom, teachers' self-efficacy meant how capable they were of drawing upon certain skills to positively influence student outcomes. I refer to these as problem-based teaching efficacy items. Problem-based teaching self-efficacy in the TSES presented tasks that teachers seemed to believe implicitly could be solved, i.e., *calming a disruptive student* and *providing an alternate explanation when students are confused*. For these items, teachers more consistently evaluated their skills in calming students or resolving student confusion, using the effectiveness of their strategies on student outcomes as relevant criteria to judge their confidence and skill. Teacher responses to these problem-based teaching tasks, therefore, seem to align with a sense of efficacy that might be highly meaningful in terms of considering teachers belief to "execute course of action" to positively influence students (Bandura, 1997; Skaalvik & Skaalvik, 2010; Tschannen-Moran & McMaster, 2009), with the courses of action being particular strategies or skills that are known to be effective for student learning. Further theoretical

support for problem-based teaching efficacy items are that inclusion of obstacles to performing the task (e.g., presenting a problem) are recommended because they provide "gradations of challenges or impediments" that prevent ceiling effects (Bandura, 2006). Indeed, teachers in this sample were only likely to rate themselves one or two points lower for items specifying circumstances in which they felt less efficacious. While only two items on the TSES were problem-based, it was encouraging that similarities were seen across teachers in the meanings of self-efficacy for these tasks, even if teachers evaluated their TSE in completely different ways with other tasks (e.g., Naomi and Florence).

Accordingly, more investigations are needed into the design, implications and meanings of problem-based teaching efficacy. One specific area of focus is defining appropriate gradations of challenges and salient problems of practice for both beginning and experienced teachers. Most teachers felt highly efficacious to provide alternate explanations when students are confused. Is this because this is less difficult problem to solve or because of self-selection into the field of teaching by those who feel confident they can clearly explain things to students? A subsequent area of research is whether self-efficacy for solving problems is strongly related to burnout and satisfaction as other forms of TSE (Skaalvik & Skaalvik, 2010) or if it holds unique predictive validity. While providing conceptual and measurement clarity is undoubtedly important for research, in tandem researchers should also consider whether greater conceptual and measurement clarity provides new information that helps untangle this "elusive construct" and its implications for both teacher and student outcomes.

The variety of interpretations that teachers brought to ISE items was concerning for reasons of validity and reliability but also because many teachers' sense of efficacy was

based on beliefs about teaching and learning not supported by empirical evidence; that is, belief that students have a particular learning style. This was most often seen in response to *TSE for Alternative Strategies* in which teachers thought about differentiating or varying instruction based on students' perceived learning styles. While variation in use of strategies is important for a variety of reasons not related to learning styles (Corno, 2008; Willingham et al., 2015), teachers' belief that this drives their use of instruction is concerning because of the attributions that teachers might make when they fail to reach students and the messages they might send to students about their capabilities to learn in any situation. Moreover, teachers who endorse a learning styles theory—which imply that a student learns best by one primary method of receiving information—may be less likely to utilize cooperative learning strategies or heterogeneous grouping because they feel students won't be able to learn from other students. This is concerning because these are strategies known to be effective for student learning in classrooms (Blanke, 2009; Corno, 2008; Kazemi, 1998). Accordingly, ongoing teacher professional development that surfaces and addresses misconceptions teachers might have about teaching and learning and why certain strategies are effective is crucial.

Just as some teachers were likely to think of learning styles as a stable attribute of students, some teachers also had a stable view of what constitutes a good question in any instructional context: high-level thinking questions. While high-level thinking questions are absolutely good questions that often are underemphasized in American classrooms (Rowan, Correnti, & Miller, 2002; Smith, Lee, & Newmann, 2001), it is also true that they might not be good questions in every single situation. Teachers who thought of good questions as those which elicit the desired purpose showed a responsiveness to students

and focus on formative assessment that is characteristic of effective instruction (Ball & Forzani, 2010; Corno, 2008). Given the variation and confusion among teachers about what is a good question, it is important for the validity and reliability of this item to clarify what is meant by "good questions." Such investigation should be guided by recent research into which types of questions or alternative strategies matter most for student learning, teacher learning, and teachers' ability to adjust their instruction to student learning.

This study has shown some preliminary evidence that the meaning of teachers' self-efficacy ratings, not just the strength, may change over the course of their careers. This was most evident in looking at experienced teachers' efficacy beliefs. While failure to achieve student outcomes (e.g., control disruptive behavior or motivate students) did not seriously undermine other teachers' self-efficacy beliefs as they believed they could learn or had improved at the task, more experienced teachers like Dennis, Amanda and Katy tended to have higher standards for success. Unlike other teacher, their lower efficacy scores reflected a desire to reach every single student or motivate students for their content and other content areas indicating a possible reference bias. This finding suggests that while these low self-efficacy ratings accurately reflect how successful experienced teachers *feel*, it may not reflect how successful they are to outside observers. However, it must also be noted that very small variations in numerical efficacy scores were seen between tasks for which teachers felt efficacious and inefficacious. Teachers tended to drop their self-efficacy score by a single point when considering that they lacked confidence or effectiveness on a particular task. This suggests that using rating scales that span a larger numerical scale, such as a 100 point scale suggested by Bandura (2006), may allow us to see potentially significant gradations in teachers' sense of efficacy. Moreover, it is unclear

that experienced teachers with somewhat lower self-efficacy in this sample were dissatisfied with their jobs to the extent that they may burnout or leave the profession, outcomes typically associated with low self-efficacy. However, since self-efficacy beliefs among teachers with over twenty years of experience were not well-represented in this sample, future studies should investigate the meaning of low efficacy beliefs of highly experienced teachers to understand if the meaning of those judgments reflects harsher standards for themselves, maladaptive beliefs about students, or actual ineffectiveness in influencing student outcomes.

Experienced teachers also revealed that important, research-based strategies to managing the learning environment may not be well-represented or solicited by current TSES items. While Dennis and Amanda spontaneously considered how teachers "set the tone" in the classroom to proactively control disruptive behavior, few other teachers considered this proactive strategy. Some teachers thought of ways to react to disruptive behavior, acknowledging that those strategies weren't highly effective in handling disruptive behavior on a classroom scale. Thus, those teachers' self-efficacy belief for this task seemed to be calibrated to their effectiveness and knowledge about effective strategies by which to achieve the outcome. This type of efficacy belief seems most similar to the problem-based teaching items; however, since other teachers' responses to this item reflected thinking about the controllability of the outcome and not the strategies by which they used to achieve the outcome, the reliability of this items could be improved. Cognitive pre-testing different forms of this item, and others, is needed in order to ascertain that the item reliably prompts teachers to consider proactive measures to control disruptive behavior through managing the environment.

Student teachers' self-efficacy beliefs in this sample can be characterized as hopefulness and conviction in the power of teachers' role to influence all students. Self-efficacy beliefs of student teachers were more likely to be influenced by successful attainment of student outcomes rather than student failure experiences, consistent with prior research (Guskey, 1987). As self-efficacy judgments are a mix of judgments of ability and potential to learn or perform a task (Usher, 2016), which means that student teachers more often considered their potential to learn, rather than current ability, given the belief that teachers can reach all students. They did not exhibit awareness or acknowledge that, despite their best efforts, they may not reach a select few students every year, unlike other teachers. It is outside the scope of this study to say whether these hopeful efficacy beliefs are healthy, or adaptive, for pre-service teachers at this point in their teaching careers, but this may potentially explain why empirical studies with the TSES often find that a one-factor solution for the multi-dimensional TSES is more appropriate for pre-service teachers than a 3-factor solution (Tschannen-Moran & Hoy, 2001). It was troublesome, however, that some pre-service teachers' efficacy beliefs were untested, such as Samantha who did not have any experiences of failure or struggle that challenged her sense of efficacy for classroom management. Challenges present opportunities for learning, growth and self-reflection if they are structured and perceived that way (Dweck, 2002). Untested self-efficacy beliefs may be more unstable, subject to precipitous drops in self-efficacy when novices enter the classroom and encounter many new challenges. Therefore, providing pre-service teachers with moderate challenges within their Zone of Proximal Development (Vygotsky & Cole, 1978) along with guidance to interpret failures as opportunities to grow, may be especially important not only for their skill development

(Daniel, Auhl, & Hastings, 2013; Valencia, Martin, Place, & Grossman, 2009) but for their resilience and stability of their efficacy beliefs.

Conclusion

In summary, the meaning of teachers' self-efficacy beliefs vary considerably depending on whether or not they focus on student outcomes, which in turn seems to depend both on item construction and the tasks under consideration. Frequently, teachers do not spontaneously consider research-based strategies known to be effective for student learning or rely on theories of learning not supported by empirical research. More consistency in the meaning of TSE was found for items that were structured as problems that needed to be solved in the classroom. Moreover, this type of TSE was a reflection upon teachers' confidence in their skills to positively influence students towards desired outcomes, not the controllability of the outcome. Altogether, this seems to suggest that problem-based teaching efficacy is an important area of future research, and that the TSES may need revisions for clarity to improve its reliability and validity.

Table 14.

Demographic information

	"Novice" Teachers (n=14)			**"Experienced" Teachers (n=9)**		**Total sample (%)**
	Pre-service teachers	**Beginning teachers**				
Length of time teaching	0-4 months	1 year	2-3 years	4-9 years	12+ years	
N	4	4	6	6	3	23 (100)
Pseudonyms	Vivian Stacy [a] Samantha [a] Darren	Bob Hannah Justin Victoria	Brianna Crystal Greg Naomi Patricia Marcus	Ally [a] Amanda [a] Caleb Florence Natalia Katy	Dennis Jennifer Paula	
n Female	3 (75%)	2 (50%)	4 (66%)	5 (83%)	2 (66%)	16 (70)
n White	3 (75%)	2 (50%)	4 (66%)	5 (83%)	1 (33%)	16 (70)
n teaching 80% or more of students eligible for free or reduced lunch		1 (25%)	4 (66%)	5 (83%)	0	11 (47)
n teaching 80% or more of students performing below grade level		2 (50%)	3 (50%)	3 (50%)	0	8 (35)

[a] Pseudonyms starting with the same letter were teaching in the same school or were from the same teacher preparation program

Table 15.

Selected codes

Code	Description of code	Sample quotations	% of units
Confused by Question	Teacher expresses confusion or is unsure what an item or part of an item means.	*'How well can you implement alternative strategies?' Alternative's totally vague. Alternative to what? To what I do or to what I'm supposed to do?*	10
SO: included	Teacher evaluates how students, in general or for their specific students, respond to the enactment of the teaching task.	*Everyone's almost always understanding what we're doing because of student explanations.*	55
SO: unclear / not mentioned	Teacher does not clearly reference how students respond to the teaching task.	*Well, I can use a great deal [of assessments]. Yes, 9, I can.*	43
SO: excluded	Teacher explicitly excludes student outcomes during efficacy evaluation because of how they interpret item intent.	*I feel like this [item] is asking to what extent I do it. It doesn't actually ask if it's effective. Because I feel like that's a different question.*	2
Inventory	Teacher response indicates a quantitative appraisal of "how much," a range of options, or how many strategies available to enact a teaching task.	*There are always a wide multitude of options [to calm a disruptive or noisy student] … Because there is that wide array, I would say it's an 8 … there are so many different options.*	20
Limitation: None / Unclear	Teacher does not clearly reference any limitation on capability to enact the teaching task.	*I would definitely say that that is definitely a strength that I personally have, I would probably say that that would be a "9."*	39
Limitation: Personal Capabilities	Teacher references not having capabilities or wanting to improve on their ability to enact the teaching task.	*When it comes to things such as grammar, I know I have a lot to work on ... we just recently did a grammar lesson, and I feel like I was more confused at the end.*	39
Limitation: Students	Teacher references student capabilities, skills, knowledge, or behavior as a limitation on their capability to enact the teaching task. This is typically in reference to certain group or type of student.	*We do have kids who have lots of issues and when they are bent on being disruptive, it will be disruptive. There's very little that we can do that's successful in bringing them back.*	12
Limitation: Teacher Context	Teacher references feeling limited by a structure or characteristic of their school or district. Could be lack of resources, time or support.	*I think if I had more time that I would definitely reach out to more families. Last year I was at a different school and there were longer block periods [and]… it gave me much more time to make contact with families.*	13
Limitation: Student context	Teacher references students' home environments, parents, or neighborhood contexts as limitations to achieving the teaching task successfully.	*There are students who have a lot of issues going on at home, that have nothing to do with the classroom … and there's very little that we can do that's successful in bringing them back.*	13
Limitation: Student responsibility	Teacher references a need for student cooperation, willingness, or responsibility in order to achieve success in the teaching task. This is not specific to a group of students but to all students in general.	*You got to throw it in the kids' court. Because you can motivate somebody forever, but if they don't do anything with what you said, or what you recommended, then they're still where they started.*	1

Table 16

Multiple meanings of TSE for student outcomes

				Patterns within theme by years teaching experience			Patterns within theme by context
Category		**Example SE statement**	**SE score influenced by consideration of student outcomes?**	**Student Teachers**	**Beginning teachers**	**Experienced teachers**	**Urban teachers**
Controllability of outcome	**Adaptive**	I can have a great deal of influence over what happens in my classroom, but not outside.	No	Teachers can and should do a lot	I can greatly influence, but not control student outcomes	I can greatly influence, but not control student outcomes. Shared responsibility with student (Stacy*)	Tended to talk about lack of support in home environment as negative influence on sense of efficacy. Only Naomi saw external environment as resource; high level of school support
	Maladaptive (experienced teachers only)	I (and others) cannot influence some students because of student environment factors. I can do (know of) lots of things to do, however.	No	n/a	n/a	External environment as negative influence on management and motivation	
Effectiveness with all students (experienced teachers only)		I have reached all students except one or two each year.	Yes	n/a	n/a	Lowered efficacy score 1-2 points when focused on one or two students per year	

					that did not reach	
Effectiveness of strategies in influencing student outcomes	I can use these strategies, and they are effective.	Yes	Strategies to react to misbehavior, rather than prevent it; little thought of classroom climate; did not identify purpose for instructional strategies or groups of students in classroom	Strategies to react to misbehavior, rather than prevent it; little thought of classroom climate; Cannot execute strategies well; Can identify student groups that they struggle to reach (Ally*)	Based on principles of management , rather than discrete strategies (Crystal*)	Thought of angry or emotional disruptive behavior (e.g., throwing pens, getting angry, being upset) that needed "calming" rather than low-level disruptive behaviors.
Ability to solve problems using particular strategies	I can solve issues that prevent learning by using pedagogical and/or content knowledge and knowledge of student(s).	Yes	Consistent meaning: I can clear up student confusion & calm a disruptive student through using my own knowledge of content and knowledge of my students.			Those with more than one year of experience (with exception of Florence*) talked about diagnosing reasons for disruptive behavior as important

* Teachers who did not fall in this category of teaching experience, but displayed thinking similar to teachers in this category

Table 17

Multiple meanings of TSE for strategies

			Patterns within theme by years teaching experience			Patterns within theme by context
Category	**Example self-efficacy statement**	**Effectiveness influence on SE rating?**	**Student Teachers**	**Beginning teachers**	**Experienced teachers**	**Urban teachers**
Variety of strategies known and/or used	I know many strategies I could use (and for some, that I do use). I can vary my instruction and assessment.	No	Did not name purpose for using strategies or groups of students who would benefit from use of strategies. Learning styles theory endorsement. Student enjoyment and interest as purpose for using strategy.	Named purposes and groups of students trying to reach with different strategies (e.g., ELL, low achieving students). Learning styles theory endorsement.		English Language Learners as subgroup (n=1)
Meet standards for performance: decontextualized	I can plan and implement instruction that aligns with external standard (e.g., Bloom's Taxonomy) in any context.	No	Bloom's Taxonomy	Bloom's Taxonomy, getting lower level students up a struggle		
Meet standards for performance: contextualized	I can plan and implement instruction that meets content goals and student needs in the moment	Yes	Consistent meaning: good questions achieve intended instructional purpose, e.g. assessment, diagnosis of student thinking/learning, or stimulate discussion.			n/a

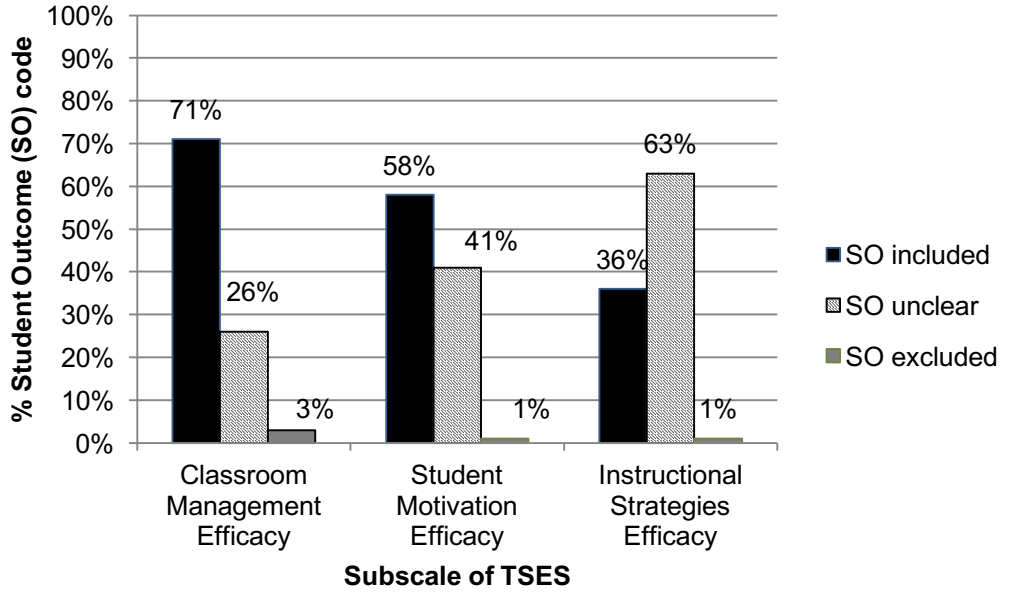

Figure 8. Application of SO code across subscales

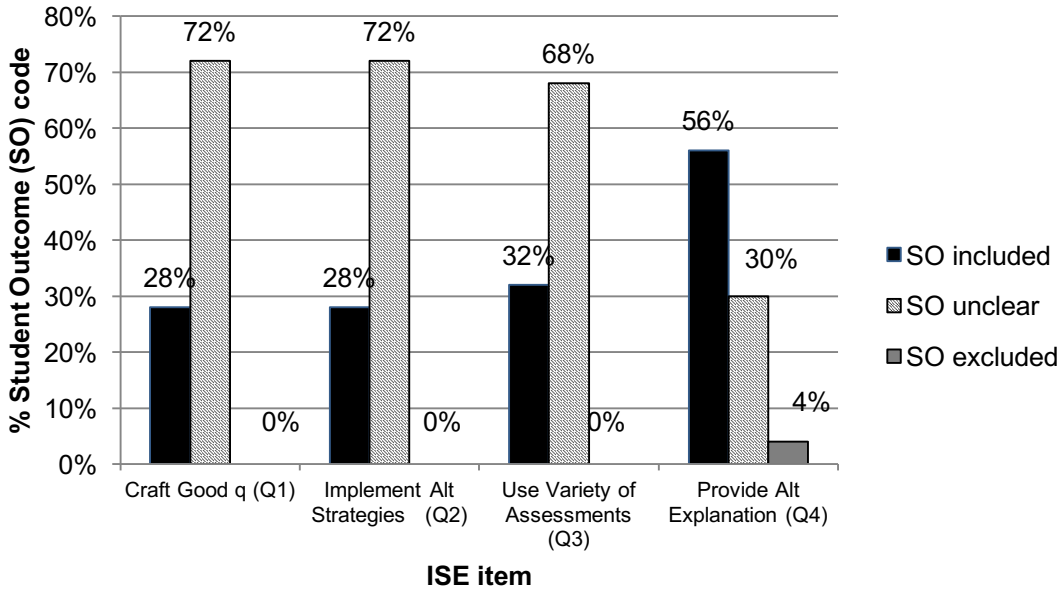

Figure 9. Application of SO code to four Instructional Strategies Efficacy (ISE) items

Appendix A: TSES short form (version 1 of 3)

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

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

Practice item: How much can you do to improve the understanding of a student who is failing?

	Response options 1 Not at all 2 3 Very little 4 5 Some degree 6 7 Quite a bit 8 9 A great deal
1. How much can you use a variety of assessment strategies?	1 2 3 4 5 6 7 8 9
2. To what extent can you provide an alternative explanation or example when students are confused?	1 2 3 4 5 6 7 8 9
3. How much can you assist families in helping their children to do well in school?	1 2 3 4 5 6 7 8 9
4. How well can you implement alternative strategies in your classroom?	1 2 3 4 5 6 7 8 9
5. How much can you do to control disruptive behavior in the classroom?	1 2 3 4 5 6 7 8 9
6. How much can you do to motivate students who show low interest in school work?	1 2 3 4 5 6 7 8 9
7. How much can you do to get students to believe they can do well in school work?	1 2 3 4 5 6 7 8 9
8. How much can you do to help your students value learning?	1 2 3 4 5 6 7 8 9
9. To what extent can you craft good questions for your students?	1 2 3 4 5 6 7 8 9
10. How much can you do to get children to follow classroom rules?	1 2 3 4 5 6 7 8 9
11. How much can you do to calm a student who is disruptive or noisy?	1 2 3 4 5 6 7 8 9
12. How well can you establish a classroom management system with each group of students?	1 2 3 4 5 6 7 8 9

APPENDIX B: Interview Protocol

Introduction

Thank you for agreeing to speak with me today! Let me start by giving you some background about the project. Often researchers ask teachers to respond to surveys or questionnaires about teaching, but we don't know what they are thinking about as they respond to the items on the survey. I am interested in how teachers interpret survey items about their teaching and what kinds of things they think about as they are trying to respond to those items. Today I have brought with me a short 12 item portion of a larger survey.

As you will see in a minute, the survey will ask that you rate on a scale of 1 to 9 "how much you can do" about certain issues related to your current teaching position. Let me explain what I'd like you to do. For each of the 12 questions or statements, I'd like you to read each aloud and then talk out loud all the thoughts that go through your mind from the moment you read the question until you make your final decision on a numerical answer and circle the number you decide reflects your opinion. Just to be clear, I'd like to hear all the thoughts that pass through your mind, no matter how trivial or "off topic" they may seem. I'm much less interested in the number you choose than I am in your thoughts as you arrive at your choice. I won't interrupt you to ask questions as you are speaking; I will wait until after you've completed the survey to ask you some follow up questions.

Let's look at the survey now and then go through the practice question. Each question asks you to respond on a scale of 1 to 9 with 1 meaning "Nothing" and 9 meaning "A great deal" [show scale at same time]. As the directions state, "please respond to each of the questions by considering the combination of your current ability, resources, and opportunity to do each of the following in your present position."

Just to make sure I've explained it clearly, we will try this thinking aloud process with a practice question. Any questions before we begin?

Practice Item:

One thing to note is that no real names will be used. If you happen to mention the name of a student, we will change that name in any reporting we do. Also, we stopped after one item, but when you are completing the survey, you'll just continue with the next question after you finish one.

This interview will take approximately 30 minutes to complete. If at any time you need to stop or take a break, just let me know. In order to make sure I capture our conversation accurately, I will be recording this conversation.

Section A: Think Aloud

[begin recording, state date and interview number]

Directions: Please read the first item and think aloud all of the thoughts that enter your mind after you read the question and while you chose your answer.

If teacher pauses for longer than 10 seconds, researcher will prompt with, "Keep talking."

Section B: Follow up questions/Cognitive Interviewing

Thank you for completing this survey. Bear with me while I ask you a few questions so I can make sure that I understand what you said in the context of the question.

1. (clarification questions for items)
   POSSIBLE FOLLOW UP for elaboration:
   - What would it mean if you had chosen a [#] instead of a [#]?
   - Can you tell me a little more about why you chose number X?

2. If not explained: Item number X asks about good questions. Can you tell me more about what the phrase good questions means?
3. Item number X asks "How well can you implement alternative strategies in your classroom?" What is that question trying to find out from you?"
4. One of the goals of this study is to see how well this survey measures what it is intended to measure. Before I tell you what that is, can you tell me what you think this survey is intended to measure?
5. This survey is intended to measure's a teacher's sense of self-efficacy. Self efficacy refers to your beliefs, or judgment, of your capability execute the courses of action required as a teacher. It's related to your confidence in your capabilities as a teacher. Knowing this, is there anything you felt you should have been asked about but weren't on this survey that contributes to your sense of self-efficacy as a teacher?
6. When or in what situations are you likely to think about your confidence in your abilities as a teacher? ("Can you tell me more about these times" if he/she hasn't mentioned frequency. If still doesn't mention frequency, "How often do these situations come up?")
7. What, if any, specific skills did you learn in your teacher preparation program that you are currently using as a teacher?

References

Ball, D. L., & Forzani, F. M. (2010). Teaching skillful teaching. *Educational Leadership*, *68*(4), 40–45.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY US: W H Freeman/Times Books/ Henry Holt & Co.

Bandura, A. (2006). Guide for Constructing Self-Efficacy Scales. In *Self-Efficacy Beliefs of Adolescents* (pp. 307–337). Information Age Publishing.

Blanke, B. L. (2009). *Understanding mathematical discourse in the elementary classroom: A case study*. ProQuest Information & Learning, US.

Borko, H., & Livingston, C. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Educational Research Journal*, *26*(4), 473–498. http://doi.org/10.2307/1162861

Brouwers, A., & Tomic, W. (2000). A Longitudinal Study of Teacher Burnout and Perceived Self-Efficacy in Classroom Management. *Teaching and Teacher Education*, *16*(2), 239–53.

Cantrell, S. C., & Callaway, P. (2008). High and low implementers of content literacy instruction: Portraits of teacher efficacy. *Teaching and Teacher Education*, *24*(7), 1739–1750. http://doi.org/10.1016/j.tate.2008.02.020

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*(6), 473–490. http://doi.org/10.1016/j.jsp.2006.09.001

Chang, M.-L. (2013). Toward a theoretical model to understand teacher emotions and teacher burnout in the context of student misbehavior: Appraisal, regulation and coping. *Motivation and Emotion*, *37*(4), 799–817. http://doi.org/10.1007/s11031-012-9335-0

Coffey, J. E., Hammer, D., & Levin, D. M. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education*, *60*(2), 142+.

Cohen, D. K. (2011). *Teaching and its predicaments*. Harvard Univ Pr.

Cohen, S. A., Hyman, J. S., Ashcroft, L., & Loveless, D. (1989). Comparing effects of metacognition, learning styles, and human attributes with alignment. In *Annual meeting of the American Educational Research Association*. San Francisco, CA.

Coladarci, T., & Breton, W. A. (1991). *Teacher Efficacy, Supervision, and the Special Education Resource-Room Teacher.* Retrieved from http://search.proquest.com.proxy.lib.umich.edu/eric/docview/63000717/13E1FA187AD44263936/1120?accountid=14667

Corbin, J. M., & Strauss, A. L. (2008). *Basics of qualitative research techniques and procedures for developing grounded theory.* Los Angeles, [Calif.] ; London: SAGE. Retrieved from http://srmo.sagepub.com/view/basics-of-qualitative-research/SAGE.xml

Corno, L. (2008). On Teaching Adaptively. *Educational Psychologist*, *43*(3), 161–173. http://doi.org/10.1080/00461520802178466

Daniel, G. R., Auhl, G., & Hastings, W. (2013). Collaborative feedback and reflection for professional growth: Preparing first-year pre-service teachers for participation in

the community of practice. *Asia-Pacific Journal of Teacher Education*, *41*(2), 159–172. http://doi.org/10.1080/1359866X.2013.777025

Domsch, G. D. (2009). *A study investigating relationships between elementary principals' and teachers' self-efficacy and student achievement*. Saint Louis University. Retrieved from http://search.proquest.com.proxy.lib.umich.edu/pqdt/docview/304997929/fulltextPDF/133E7064D425DA0A64/1?accountid=14667

Donald, K. L. (2009). *Evaluation of self-reported teacher efficacy and minority achievement in middle school*. ProQuest Information & Learning, US.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*(4), 237–251. http://doi.org/10.3102/0013189X15584327

Duffin, L. C., French, B. F., & Patrick, H. (2012). The Teachers' Sense of Efficacy Scale: Confirming the factor structure with beginning pre-service teachers. *Teaching and Teacher Education*, *28*(6), 827–834. http://doi.org/10.1016/j.tate.2012.03.004

Dweck, C. S. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways). In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education.* (pp. 37–60). San Diego, CA US: Academic Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. the MIT Press.

Feldon, D. F. (2007a). Cognitive Load and Classroom Teaching: The Double-Edged Sword of Automaticity. *Educational Psychologist*, *42*(3), 123–137. http://doi.org/10.1080/00461520701416173

Feldon, D. F. (2007b). The implications of research on expertise for curriculum and pedagogy. *Educational Psychology Review*, *19*(2), 91–110. http://doi.org/10.1007/s10648-006-9009-0

Fenstermacher, G. D., & Richardson, V. (2005). On Making Determinations of Quality in Teaching. *Teachers College Record*, *107*(1), 186–213. http://doi.org/10.1111/j.1467-9620.2005.00462.x

Gabriele, A. J., & Joram, E. (2007). Teachers' reflections on their reform-based teaching in mathematics: Implications for the development of teacher self-efficacy. *Action in Teacher Education*, *29*(3), 60–74.

Gonzalez, N., Moll, L. C., Tenery, M. F., Rivera, A., Rendon, P., Gonzales, R., & Amanti, C. (1995). Funds of Knowledge for Teaching in Latino Households. *Urban Education*, *29*(4), 443–470. http://doi.org/10.1177/0042085995029004005

Grossman, P., Compton, C., Shahan, E., Ronfeldt, M., Igra, D., & Shaing, J. (2007). Preparing practitioners to respond to resistance: a cross-professional view. *Teachers and Teaching: Theory and Practice*, *13*(2), 109–123.

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

Henson, R. K. (2001). Teacher Self-Efficacy: Substantive Implications and Measurement Dilemmas. Retrieved from

http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED452208

Jackson, P. W. (1990). *Life in Classrooms*. Teachers College Press.

Jacobs, V. R., Lamb, L. L. C., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, *41*(2), 169–202.

Karabenick, S. A., Woolley, M. E., Friedel, J. M., Ammon, B. V., Blazevski, J., Bonney, C. R., … others. (2007). Cognitive Processing of Self-Report Items in Educational Research: Do They Think What We Mean? *Educational Psychologist*, *42*(3), 139–151.

Kazemi, E. (1998). Discourse That Promotes Conceptual Understanding. *Teaching Children Mathematics*, *4*(7), 410–14.

Kim, Y. H., & Kim, Y. E. (2010). Korean early childhood educators' multi-dimensional teacher self-efficacy and ECE center climate and depression severity in teachers as contributing factors. *Teaching and Teacher Education*, *26*(5), 1117–1123. http://doi.org/10.1016/j.tate.2009.06.009

Klassen, R. M., Durksen, T. L., & Tze, V. M. C. (2014). Teachers' self-efficacy beliefs: Ready to move from theory to practice? In P. W. Richardson, S. A. Karabenick, & H. M. G. Watt (Eds.), *Teacher motivation: Implications for theory and practice* (pp. 100–115). U.S.: Routledge/Taylor & Francis.

Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, *12*, 59–76. http://doi.org/10.1016/j.edurev.2014.06.001

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher Efficacy Research 1998–2009: Signs of Progress or Unfulfilled Promise? *Educational Psychology Review*, *23*(1), 21–43. http://doi.org/10.1007/s10648-010-9141-8

Kleinsasser, R. C. (2014). Teacher efficacy in Teaching and Teacher Education. *Teaching and Teacher Education*, *44*, 168–179. http://doi.org/10.1016/j.tate.2014.07.007

Klieme, E. (2012, August). *Qualities and effects of teaching: Towards a conceptual theory of teaching.* Presented at the EARLI SIG meeting, Educational Effectiveness, Zurich, Switzerland.

Korthagen, F., Loughran, J., & Russell, T. (2006). Developing fundamental principles for teacher education programs and practices. *Teaching and Teacher Education*, *22*(8), 1020–1041. http://doi.org/10.1016/j.tate.2006.04.022

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. http://doi.org/10.1207/s15430421tip4104_2

Labone, E. (2004). Teacher Efficacy: Maturing the Construct through Research in Alternative Paradigms. *Teaching & Teacher Education*, *20*(4), 341–359.

Lampert, M., Boerst, T. A., & Graziani, F. (2011). Organizational Resources in the Service of School-Wide Ambitious Teaching Practice. *Teachers College Record*, *113*(7), 1361–1400.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook*. Thousand Oaks: Sage Publications. Retrieved from http://hdl.handle.net/2027/[u]: mdp.39015050352072

Moè, A., Pazzaglia, F., & Ronconi, L. (2010). When being able is not enough. The combined value of positive affect and self-efficacy for job satisfaction in teaching.

Teaching and Teacher Education, *26*(5), 1145–1153. http://doi.org/10.1016/j.tate.2010.02.010

Mulholland, J., & Wallace, J. (2001). Teacher Induction and Elementary Science Teaching: Enhancing Self-Efficacy. *Teaching and Teacher Education*, *17*(2), 243–61.

Pianta, R. C., Belsky, J., Vandergrift, N., Houts, R., & Morrison, F. J. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal*, *45*(2), 365–397. http://doi.org/10.3102/0002831207308230

Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education*, *65*(2), 150–167. http://doi.org/10.2307/2112680

Ross, J. A., Hogaboam-Gray, A., & Hannay, L. (2001). Effects of teacher efficacy on computer skills and computer cognitions of Canadian students in Grades K–3. *The Elementary School Journal*, *102*(2), 141–156. http://doi.org/10.1086/499697

Ross, J. A., McKeiver, S., & Hogaboam-Gray, A. (1997). Fluctuations in Teacher Efficacy During Implementation of Destreaming. *Canadian Journal of Education*, *22*(3), 283–96.

Rowan, B., Correnti, R., & Miller, R. (2002). What Large-Scale Survey Research Tells Us About Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools. *The Teachers College Record*, *104*(8), 1525–1567.

Ryan, A. M., Kuusinen, C. M., & Bedoya-Skoog, A. (2015). Managing peer relations: A dimension of teacher self-efficacy that varies between elementary and middle

school teachers and is associated with observed classroom quality. *Contemporary Educational Psychology*, *41*, 147–156. http://doi.org/10.1016/j.cedpsych.2015.01.002

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*(3), 611–625. http://doi.org/10.1037/0022-0663.99.3.611

Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, *26*(4), 1059–1069. http://doi.org/10.1016/j.tate.2009.11.001

Skinner, E. A. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology*, *71*(3), 549–570. http://doi.org/10.1037/0022-3514.71.3.549

Smith, J. B., Lee, V. E., & Newmann, F. M. (2001). *Instruction and Achievement in Chicago Elementary Schools. Improving Chicago's Schools.* (p. 51). Retrieved from http://search.proquest.com.proxy.lib.umich.edu/eric/docview/62190775/13F3DF0B6B32F6763F/1?accountid=14667

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. http://doi.org/10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*(2), 202–248. http://doi.org/10.2307/1170754

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. http://doi.org/10.1086/605771

Tsigilis, N., Koustelios, A., & Grammatikopoulos, V. (2010). Psychometric properties of the Teachers' Sense of Efficacy Scale within the Greek educational context. *Journal of Psychoeducational Assessment*, *28*(2), 153–162. http://doi.org/10.1177/0734282909342532

Usher, E. L. (2016). Personal capability beliefs. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (3rd ed., pp. 146–159). Routledge.

Valencia, S. W., Martin, S. D., Place, N. A., & Grossman, P. (2009). Complex Interactions in Student Teaching: Lost Opportunities for Learning. *Journal of Teacher Education*, *60*.

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: the development of higher psychological processes*. Cambridge: Harvard University Press.

Wang, H., Hall, N. C., & Rahimi, S. (2015). Self-efficacy and causal attributions in teachers: Effects on burnout, job satisfaction, illness, and quitting intentions. *Teaching and Teacher Education*, *47*, 120–130. http://doi.org/10.1016/j.tate.2014.12.005

Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education*, *21*(7), 747–766. http://doi.org/10.1016/j.tate.2005.05.009

Willingham, D. T., Hughes, E. M., & Dobolyi, D. G. (2015). The scientific status of learning styles theories. *Teaching of Psychology*, *42*(3), 266–271. http://doi.org/10.1177/0098628315589505

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

As discussed in the previous chapters, teacher self-efficacy (TSE) is the most widely researched teacher belief that has shown strong associations with teacher satisfaction, intent to stay in the field, and burnout (Brouwers & Tomic, 2000; Skaalvik & Skaalvik, 2007). Despite relatively recent advances in the conceptualization and measurement of TSE (Tschannen-Moran & Hoy, 2001), much less is known about how strongly TSE is associated with research-based practices known to be effective for student learning and the meaning of these task-specific beliefs in terms of teachers' thinking about effectiveness, with particular attention to student outcomes. The purpose of this dissertation, therefore, was to provide need insight into the relationship between TSE and effective teaching practices from a synthesis the extant literature and from the perspective of current middle and high school teachers. Within each of the two research studies presented in Chapters 3 and 4, I discussed implications and directions for future research. In this chapter, I elaborate these implications to discuss a) conceptualizing and measuring TSE as teaching self-efficacy for solving problems of practice and b) a consideration of the context as applied to student teachers. In introduce the term "teaching self-efficacy" in this chapter to refer to a focus on self-efficacy for solving problems of practice.

Problems of practice approach to conceptualizing and measuring teaching self-efficacy

While TSE has been reconceptualized to align more strongly with a task-based social cognitive theory of self-efficacy, the field has not sufficiently and comprehensively addressed the tasks of teaching that form the basis of this important belief. Recent theoretical advances within teaching and teacher education have more clearly defined different domains of effective instruction and the observable and measurable practices that represent effective teaching within each domain (Brophy & Good, 1986; Klieme, 2012; Pianta & Hamre, 2009). As reviewed in Chapters 2 and 3, the teaching tasks that comprise multi-dimensional TSE measures do not represent the breadth nor the depth of tasks within corresponding domains of effective classroom instruction. Moreover, teachers studied in Chapter 4 did not spontaneously consider such research-based practices when they responded to items, at times reflecting their beliefs about teaching and learning that have no empirical support.

Moreover, teachers' efficacy beliefs took on different meanings: controllability of the outcome rather than effectiveness with students, knowledge of strategies to achieve the outcome, ability to execute research-based strategies with skill, effectiveness with every single student, or meeting external standards for performance. *Given the many different ideas about what it means to feel confident to "execute the courses of action" in teaching, the only common baseline of TSE seen in Chapter 4 seems to be how confident and competent they feel they can teach, however that was defined.* Teachers' answers reflected some consideration of how effective they were in influencing students, some consideration of the extent of their professional knowledge, some consideration of their willingness to keep trying, and some consideration of their confidence they would improve or had already improved. This belief manifested as a small but significant relationship with observers'

assessment of teaching quality in Chapter 3, with an effect size similar to that found between self-beliefs and performance in the workplace, particularly for complex tasks (e.g., Dunning, Heath & Suls, 2004).

This is a conundrum, however, since self-efficacy by definition is assumed to be a task-specific belief rather than a general self-perception of oneself (Bong & Skaalvik, 2003), and its power (and my interest in self-efficacy as applied to the teaching profession) is rooted in its power to predict task-specific behavior. The more confident teachers feel confident to do a task, the more likely they should engage in that task. If that task is one that is known to be particularly important for student learning—such as interpreting formative assessment data for the purposes of re-designing instruction or supporting ELL students to engage in academic discussions—then teacher's self-efficacy for that task should be a powerful predictor of effective teaching practice. TSE beliefs can tell us not only what a teacher knows but how confident a teacher is they can apply that knowledge to accomplish something worthy of being accomplished. By worthy, I mean a task that is important for leveraging student capacity to learn from the teacher, from other students, and from the students' own reflection on experience. By worthy, I mean research-based and culturally appropriate. At the broadest level, worthy tasks are those that have the potential to close achievement gaps, provide students with skills they need to succeed beyond the walls of the classroom, and empower all students to discover resources for their own learning and development. As a task-specific motivational belief, one could conclude that the field of TSE has failed at this point to fully consider whether we are assessing TSE for worthy teaching tasks. As a part of this critical consideration, concurrent investigation into how best to measure self-efficacy for those teaching tasks must also occur.

Findings from Chapter 4 provide preliminary evidence of an approach to assessing TSE that may hold value in terms of drawing upon teachers' capacity to apply professional knowledge for the benefit of student learning, motivation, and behavior. When teachers were prompted to consider whether they could address a specific problem such as student confusion or a disruptive and noisy student, their self-efficacy judgments more often reflected judgments of how successfully they could draw upon professional knowledge and skill to solve this problem that was inhibiting student learning. A problem-based approach to measuring teaching self-efficacy is fruitful for many reasons, three of which I highlight here. First, the nature of K-12 teaching in a compulsory education system that often recreates and reifies societal inequalities (Anyon, 1980; Powell, Farrar, & Cohen, 1985; Stanton-Salazar & Dornbusch, 1995) creates endemic problems that arise in the course of instruction. Such problems include the need to cover a wide range of topics but limited time in which to do so; managing diverse levels of background knowledge and skill while maintaining high expectations for all students; and building relationships with students who may not be inclined to trust teachers or desire to be there (Cohen, 2011; Cohen, Raudenbush, & Ball, 2003). Thus, a problem-based teaching self-efficacy evaluation should focus on the heart of what teachers do in the classroom while also providing an obstacle or challenge related to student thinking, behavior or attitude, thus providing a metric by which teachers may judge their competence (Bandura, 2006).

Second, teachers can anticipate problems with greater experience and expertise (Berliner, 2004; Borko & Livingston, 1989), yet in order to solve such problems with appropriate strategies, teachers must draw upon knowledge of their current students' skills, dispositions and attitudes. For this reason, a problem-based teaching self-efficacy

evaluation will necessarily be situated in the teachers' specific context. This may help mitigate the chance that teachers will judge their self-efficacy for tasks largely based on the controllability of the outcome, as happened with Naomi and Florence in Chapter 4, or on beliefs about what teachers should or could be doing, as was common with student teachers.

Third, a problem identified in this study and others is that teachers generally rate their sense of efficacy quite highly, thus contributing to a ceiling effect. This raises questions whether current TSE measures capture the full range of practices and difficulty levels in order to adequately differentiate between high and low efficacy teachers. In particular, in Chapter 4 we saw that high self-efficacy ratings can mask maladaptive beliefs about student performance in the face of negative environmental influences, such as with Florence. However, there was also evidence that Florence and other teachers like Naomi who had rated themselves as highly efficacious for other classroom management tasks lowered their self-efficacy ratings for a problem-based classroom management task (calming a disruptive student) because they did not feel they had the knowledge or skill to solve that problem. Supporting this finding, teachers' self-efficacy for problem solving, particularly for behavior management, is typically lower than general assessments of their self-efficacy (Chang, 2013). However, it must also be noted that the professional knowledge and skills they were drawing upon to solve this problem—relationship-building with students—actually is more in line with the domain of creating a supportive climate in their classrooms. For this reason, problem-based self-efficacy items need to be developed and tested with teachers of varying levels of experience because the kinds of knowledge teachers draw upon to solve particular problems may not be clear to researchers. Moreover,

situating the task in the daily classroom work of teachers (rather than planning to teach) calls upon different, harder to learn skills. For example, in Chapter 4 some novice teachers in this study spontaneously differentiated between preparation for teaching and the in-the-moment act of teaching students and making decisions, feeling less efficacious for the latter.

Finally, it may not be clear how common particular problems, or tasks, are generalizable across contexts. All teachers felt that they frequently needed to provide alternate explanations when students were confused (construed as a problem), and only one first-year teacher felt slightly less efficacious to do so. Calming disruptive students (construed as a problem) was a more frequent and necessary task for urban teachers to solve compared to suburban teachers in Chapter 4. Suburban teachers, on the other hand, felt that it was less necessary to involve parents to help their students do well (construed as a task) compared to urban teachers. The relevance of this task seemed to influence some teachers' judgments in distinct ways: some simply evaluated what they could do if they wanted while others rated themselves lower because they felt they should or could be doing more. In this case, a sense of responsibility (Lauermann & Karabenick, 2011) and teachers' perception of their role (Hall, Draper, Smith, & Bullough, 2008) might influence their TSE judgments when the task is not one particularly necessary or common in their teaching contexts. Although it was not common for teachers in suburban contexts, like Dennis, to calm students, he did still judge his self-efficacy in similar ways: drawing upon his knowledge to calm a student. Further investigation is needed into the kinds of problems that are both useful to be solved and that require teachers to draw upon research-based practices, skills or knowledge in order to solve them.

Student teachers and their teaching contexts

Evidence from Chapter 4 suggests that TSES items are generally successful in soliciting teachers' thinking about their current context and students. However, there was some evidence that student teachers at times relied on thinking about what is possible for teachers to do rather than what they were capable of doing given their current students. Moreover, they more often felt less efficacious because of restrictions on their responsibilities and duties as student teachers, not because of their own skills—for example, they did not have access to parent contact information or real experience interacting with parents. Finally, they were less likely to spontaneously think about the purposes for using particular strategies in the classroom compared to their more experienced colleagues.

These findings suggest reasons why studies have found low discrimination between TSE scales among pre-service teachers (Çakır & Alıcı, 2009; Tschannen-Moran & Hoy, 2007). I suggest that is because of the unique contexts in which student teachers act as teachers. Thus the need to first consider the type of self-efficacy information that is useful for researchers to know in relation to pre-service teachers' teaching *contextual* self-efficacy beliefs. Like in-service teachers, the context for student teachers is the immediate context of the classroom full of students and the school in which it is situated; unlike in-service teachers, it is also the distal context of their teacher preparation programs (Ronfeldt, Reininger, & Kwok, 2013). Thus, pre-service teachers situate their beliefs in at least two distinct, and possibly contradictory, contexts (Borko & Putnam, 1996; Middleton, Abrams, & Seaman, 2011). Pre-service teachers may learn one strategy in their teacher preparation programs, but that strategy is not enacted or supported by their cooperating

teacher (the one in charge of the classroom). Thus, the conflict of contexts may highly impact their sense of efficacy to enact that strategy or apply their professional knowledge to solve problems of practice. I can foresee two ways in which to better clarify the influence of these two contexts for student teacher TSE beliefs.

First, assessing student teachers' teaching self-efficacy beliefs as experienced in their present context given the restriction on their duties, therefore, would likely provide information about the congruence, or lack thereof, between the two contexts and its influence on TSE beliefs and current behaviors. Such TSE beliefs may be important for educators and researchers wishing to understand how students will perform at the end of the semester and whether or not they are satisfied with their chosen profession and wish to continue. Such an evaluation would be useful for intervening and supporting student teachers immediately, but not as useful for predicting how student teachers can act when they have full responsibility for student learning.

On the other hand, one might assess pre-service teachers' TSE beliefs in their present context as if they were the teachers of record. This TSE belief would provide more information regarding what student teachers feel capable of doing given their present knowledge and skills. While this may not predict behavior in that context, it may be more useful in terms of predicting what kinds of practices teachers will engage in as first-year teachers, given their experience with the students they have taught now. Such a TSE belief would be useful to compare beliefs among student teachers from the same teacher preparation program yet student teaching in different schools to determine how the immediate context impacts this type of future-oriented TSE belief. Limitations of such an assessment of TSE might be that student teachers have unrealistic expectations about what

might be possible as a teacher of record and that they still may not have experience with particular tasks, such as contacting parents, in order to make reasoned evaluations.

This consideration of relevant experience brings me back to the earlier point that a problem-based approach to assessing teaching self-efficacy must carefully consider what types of problems teachers are most likely to encounter in their present contexts. While it is certainly useful to know if student teachers feel confident to engage parents who do not speak English as a first language (a problem-based approach), given lack of any relevant experience with parents, such a teaching self-efficacy evaluation is less likely to draw upon professional teaching skills needed for that task. For example, while engaging with parents successfully requires using clear and plain language used and soliciting parents' cultural knowledge resources (Gonzalez et al., 1995), student teachers without such experience may instead evaluate their skill to solve this problem using personal characteristics such as their charm, ability to speak slowly, and comfort speaking with new people. While these characteristics may help, they may not help solve this problem at a high level, and student teachers may not be aware of it. Thus, to elicit a teaching self-efficacy evaluation that draws upon professional knowledge learned up to this point, it is crucial that significant problems of practice *in which beginning teachers have some professional knowledge and experience* are identified. To this, one can draw from work in education and teacher preparation programs across the nation to identify the kinds of knowledge and skill we want beginning teachers to be able to do upon leaving their teacher preparation programs (Ball & Forzani, 2010; Lampert et al., 2013). However, particular types of knowledge, skills and practices may be unique to certain teacher preparation programs.

Finally, as student teachers were not entirely aware of the most important purposes of using particular strategies (e.g., multiple modalities), these purposes must be made explicit for them to provide more accurate assessments and to increase the cognitive validity of the items. A well-formulated problem-based approach may be helpful in clarifying the purpose, which is to resolve the particular problem stated in the item.

In all, teaching self-efficacy conceptualized and measured as teachers' beliefs in their capabilities to solve problems of practice, holds promise as a potential new way to improve the reliability, construct and predictive validity of TSE as a task-specific judgment of capabilities. If evaluations of teaching self-efficacy consistently elicit teacher reflection on their knowledge and ability to apply appropriate effective teaching strategies for their students, we may finally be able to move TSE from theory and its assessment into practice.

References

Anyon, J. (1980). Social class and the hidden curriculum of work. *The Journal of Education*, *162*(1), 67–92.

Ball, D. L., & Forzani, F. M. (2010). Teaching skillful teaching. *Educational Leadership*, *68*(4), 40–45.

Bandura, A. (2006). Guide for Constructing Self-Efficacy Scales. In *Self-Efficacy Beliefs of Adolescents* (pp. 307–337). Information Age Publishing.

Berliner, D. C. (2004). Describing the Behavior and Documenting the Accomplishments of Expert Teachers. *Bulletin of Science, Technology & Society*, *24*(3), 200–212.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40. http://doi.org/10.1023/A:1021302408382

Borko, H., & Livingston, C. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Educational Research Journal*, *26*(4), 473–498. http://doi.org/10.2307/1162861

Borko, H., & Putnam, R. T. (1996). Learning to teach. In D. C. Berliner, R. C. Calfee, D. C. (Ed) Berliner, & R. C. (Ed) Calfee (Eds.), *Handbook of educational psychology.* (pp. 673–708). New York, NY, US; London, England: Macmillan Library Reference Usa.

Brophy, J. E., & Good, T. L. (1986). Teacher behaviors and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). New York: Macmillan.

Brouwers, A., & Tomic, W. (2000). Disruptive Student Behavior, Perceived Self-Efficacy, and Teacher Burnout. Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED450120

Çakır, Ö., & Alıcı, D. (2009). Seeing self as others see you: Variability in self-efficacy ratings in student teaching. *Teachers and Teaching: Theory and Practice*, *15*(5), 541–561. http://doi.org/10.1080/13540600903139555

Chang, M.-L. (2013). Toward a theoretical model to understand teacher emotions and teacher burnout in the context of student misbehavior: Appraisal, regulation and coping. *Motivation and Emotion*, *37*(4), 799–817. http://doi.org/10.1007/s11031-012-9335-0

Cohen, D. K. (2011). *Teaching and its predicaments*. Harvard Univ Pr.

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, Instruction, and Research. *Educational Evaluation and Policy Analysis*, *25*(2), 119–142. http://doi.org/http://dx.doi.org.proxy.lib.umich.edu/10.3102/01623737025002119

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*(3), 69–106.

Gonzalez, N., Moll, L. C., Tenery, M. F., Rivera, A., Rendon, P., Gonzales, R., & Amanti, C. (1995). Funds of Knowledge for Teaching in Latino Households. *Urban Education*, *29*(4), 443–470. http://doi.org/10.1177/0042085995029004005

Hall, K. M., Draper, R. J., Smith, L. K., & Bullough, R. V. (2008). More than a place to teach: exploring the perceptions of the roles and responsibilities of mentor teachers. *Mentoring & Tutoring: Partnership in Learning*, *16*(3), 328–345.

Klieme, E. (2012, August). *Qualities and effects of teaching: Towards a conceptual theory of teaching.* Presented at the EARLI SIG meeting, Educational Effectiveness, Zurich, Switzerland.

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., … Crowe, K. (2013). Keeping It Complex Using Rehearsals to Support Novice Teacher Learning of Ambitious Teaching. *Journal of Teacher Education*, *64*(3), 226–243. http://doi.org/10.1177/0022487112473837

Lauermann, F., & Karabenick, S. A. (2011). Taking teacher responsibility into account(ability): Explicating its multiple components and theoretical status. *Educational Psychologist*, *46*(2), 122–140. http://doi.org/10.1080/00461520.2011.558818

Middleton, M., Abrams, E., & Seaman, J. (2011). Resistance and Disidentification in Reflective Practice with Preservice Teaching Interns. *New Directions for Teaching and Learning*, (126), 67–75.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity. *Educational Researcher*, *38*(2), 109–119. http://doi.org/10.3102/0013189X09332374

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The Shopping Mall High School. Winners and Losers in the Educational Marketplace.* Retrieved from http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ326589

Ronfeldt, M., Reininger, M., & Kwok, A. (2013). Recruitment or Preparation? Investigating the Effects of Teacher Characteristics and Student Teaching. *Journal of Teacher Education*, *64*(4), 319–337. http://doi.org/10.1177/0022487113488143

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*(3), 611–625. http://doi.org/10.1037/0022-0663.99.3.611

Stanton-Salazar, R. D., & Dornbusch, S. M. (1995). Social Capital and the Reproduction of Inequality: Information Networks among Mexican-Origin High School Students. *Sociology of Education*, *68*(2), 116–135. http://doi.org/10.2307/2112778

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. http://doi.org/10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education*, *23*(6), 944–956.