

ABSTRACT

Title of Document: AN EXPERIMENTAL EVALUATION OF THE EFFECT OF INSTRUCTIONAL CONSULTATION TEAMS ON TEACHER EFFICACY: A MULTIVARIATE, MULTILEVEL EXAMINATION

Jessica Robyn Koehler, Doctor of Philosophy, 2010

Directed By: Professor Gary Gottfredson
Counseling and Personnel Services

Teacher efficacy, the extent to which teachers feel they can influence student learning (Berman, McLaughlin, Bass, Pauly, & Zellman, 1977), has been repeatedly linked to important student and teacher outcomes (Gibson & Dembo, 1984). Although the results of many studies support the claim that teacher efficacy is an important educational construct, few studies have investigated interventions to influence these teacher beliefs. The current study evaluated whether a specific teacher intervention, Instructional Consultation Teams (IC Teams), positively affected teachers' sense of self-efficacy as measured by two efficacy instruments. Participants included 1203 in-service elementary school teachers in 34 elementary schools within a large suburban school district—17 randomly assigned to the IC Team intervention and 17 assigned to the control condition. Because teachers are nested within schools, hierarchical linear modeling was utilized to

evaluate whether scores on measures of teacher self-efficacy were influenced by IC Teams. A multivariate model was also used to evaluate the effects of IC Teams on both measures, simultaneously. The results imply that IC Teams significantly increased teachers' scores on the efficacy scales. The current study provides one of a few attempts to evaluate the effects of a specific school intervention on teacher efficacy within an experimental framework.

AN EXPERIMENTAL EVALUATION OF THE EFFECT OF INSTRUCTIONAL
CONSULTATION TEAMS ON TEACHER EFFICACY: A MULTIVARIATE,
MULTILEVEL EXAMINATION

by

Jessica Robyn Koehler

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:

Professor Gary Gottfredson, Chair
Professor Sylvia Rosenfield
Associate Professor William Strein
Professor Patricia Alexander
Professor Dennis Kivlighan

Acknowledgements

The completion of my doctoral dissertation would not have been possible without the love, support, and guidance of my family, friends, and professors. To everyone who provided encouragement and guidance, I extend my warmest gratitude.

My dissertation chair, Dr. Gary Gottfredson, deserves a special thank you. Gary, your encouragement and guidance enabled me to have the confidence to undertake and complete my dissertation. Without your support, I would not have attempted to write such a complex paper. Please understand how much I appreciate all of the time you gave to help me complete my dissertation. In addition, please know how much you have influenced my future professional path.

Dr. Sylvia Rosenfield also deserves a special thank you. Sylvia, your passion for your work is truly inspiring! You re-ignited my passion for our profession. I hope you are aware of how much I appreciate your mentorship. Thank you for all of your guidance and for allowing me the opportunity to work on such a wonderful project.

I would also like to thank my other committee members, Dr. William Strein, Dr. Patricia Alexander, and Dr. Dennis Kivlighan. Your encouragement and feedback were excellent and enabled me to complete my dissertation. It was such a pleasure to work with each of you.

I would also like to thank the entire IC Teams research team (past and present), with a special thanks to Kate Bruckman and Phuong Vu. You both were very helpful in providing me with the data I needed to complete my study.

On a personal note, although I am thankful to all of my family members, two women in my family inspired me to become a school psychologist and provided

continual encouragement to complete my dissertation. My mother, Becky Savani Gilbert, a special education teacher, is truly and has always been the “wind beneath my wings.” Without your love and support, I would not be the person I am today. Thank you mom! My grandmother, Genevieve Savani, was a school psychologist beginning in the 1940’s. Grammy, thank you so much for your support and guidance and for inspiring me to become a school psychologist!

The person who is deserving of the most gratitude is my husband, Chris Koehler. Without your love and support, I never would have finished my dissertation. Your confidence in my abilities and pride in my accomplishments is truly moving. Finally, Alexander Fox, although you made the completion of my dissertation significantly more difficult, you also inspire me to be a better person. Chris and Alex, I love you both beyond words!

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iv
List of Tables.....	v
Literature Review.....	1
Teacher Efficacy: An Important Educational Construct.....	1
History of Teacher Efficacy Measurement.....	3
Interventions that Affect Teacher Efficacy.....	4
<i>Standards-Based Professional Development Program</i>	10
<i>Cognitive Coaching</i>	12
<i>Dare to be You Teacher Training</i>	15
<i>Responsive Classroom Approach</i>	16
Instructional Consultation Teams.....	18
<i>IC Teams Hypothesized Effect of Teacher Efficacy</i>	21
Consultation and Teacher Efficacy.....	22
Research Questions and Framework.....	24
Method.....	25
Subjects.....	26
Data Collection Procedures.....	26
Measures.....	27
<i>Teacher Efficacy Scales</i>	27
<i>Teacher Demographics</i>	30
Intervention Design and Implementation.....	30
Research Design.....	32
Power Analysis.....	32
Multilevel Analysis.....	33
Analytic Approach.....	33
Attrition and Treatment Diffusion.....	36
Multilevel, Multivariate Model.....	36
Results.....	40
Attrition and Treatment Diffusion.....	40
Unconditional Model.....	41
Correlations at Level-2 and Level.....	41
Within-School Model.....	42
Between-School Model.....	43
Proportion of Variance Explained.....	45
Multivariate Testing of Total Treatment Effect.....	46
Analyses with Cases with Complete Data.....	46
Discussion.....	47
Treatment Effect.....	47
Relationship Between the Teacher Efficacy Scales.....	48
Relationship Between Efficacy Scales and Predictors.....	48
Limitations.....	49
Future Research.....	51
Implications.....	53

Appendix A.....	55
Appendix B.....	56
Appendix C.....	57
Author Note.....	58
References.....	71

List of Tables

Table 1: Factor Loadings for TISA and TSES Efficacy for Instructional Strategies Scale.....	59
Table 2: Level of Implementation and Utilization for Treatment Schools.....	61
Table 3: Descriptive Statistics for TISA and TSES Efficacy for Instructional Strategies Scale.....	62
Table 4: Teacher-level Predictor Variables.....	63
Table 5: Correlations Among Teacher-Level Variables.....	64
Table 6: Correlations Between Scales at Each Level.....	65
Table 7: Final Within School Model, Grand Mean Centered with Fixed Effects.....	66
Table 8: Final Between School Model.....	67
Table 9: Variance Components and Proportion of Variance Explained.....	68
Table 10: Final Between School Model Dropping Cases with Missing Data.....	69
Table 11: Descriptive Statistics for the TISA and TSES Efficacy for Instructional Strategies.....	70

Literature Review

Research has repeatedly demonstrated a strong relation between beliefs, behavior, and outcomes. Two influential psychological theories of the twentieth century, self-efficacy and locus of control, focus on personal beliefs and their influence on human behavior and outcomes. Self-efficacy concerns whether a person believes he or she has the capability to produce a given action (Bandura, 1997), while locus of control focuses on whether outcomes are determined by external forces or one's own actions (Rotter, 1966). These concepts are similar and sometimes incorrectly viewed as the same construct. However, locus of control and self-efficacy are distinct constructs. Beliefs about whether actions influence outcomes (locus of control) are not equivalent to beliefs about whether one can produce certain actions (self-efficacy).

One line of investigation that evolved from a blending of self-efficacy and locus of control theories is teacher efficacy research (Rose & Medway, 1981). Three decades of inquiry indicate that teacher efficacy is related to various educational outcomes, such as student achievement and motivation (Tournaki & Podell, 2005; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). Defined as the extent to which a teacher believes he or she has the capacity to affect student performance (Berman, McLaughlin, Bass, Pauly, & Zelman 1977), teacher efficacy has been repeatedly linked to important educational variables.

Teacher Efficacy: An Important Educational Construct

Teacher sense of efficacy is important because it is related to student achievement. An investigation by Rand Corporation researchers (Armor et al., 1976) became the catalyst for decades of teacher efficacy research. Armor et al. found that the

greater the efficacy scores of teachers, the more their students gained on the reading portion of the California Test of Basic Skills. Later researchers, (Anderson, Greene, & Loewen, 1988; Ashton & Webb, 1986; Ross, 1992) have also provided support for the link between teacher efficacy and student achievement. Ross (1998) found that teachers scoring higher on teacher efficacy measures were more likely to attempt innovative teaching practices that led to increased student achievement. Aside from achievement, other student outcomes, such as motivation and students' own sense of self-efficacy, have been positively correlated with teacher efficacy (Anderson, Greene, & Loewen, 1988; Midgely, Fellauffer, & Eccles, 1989).

In addition to student variables, teacher efficacy has been associated with many important teacher characteristics. For example, Allinder (1994) found that teachers with higher efficacy displayed better planning and organizational skills in the classroom. Several researchers (Coldarci, 1992; Evans & Tribble, 1986; Gibson & Dembo, 1984) have shown that teachers with greater efficacy have more commitment to struggling students and refer fewer students to special education (Soodak & Podell, 1993). Teacher efficacy has been found to be related to success in curriculum innovation (Berman et al., 1977), quality of student-teacher relationships (Ashton & Webb, 1986), confidence in working with parents (Hoover-Dempsey, Bassler, & Brissie, 1987), time spent on academic learning (Allinder, 1995), ability to hold students accountable for their own learning (Ashton & Webb, 1986), teacher valuing of educational innovations (Cousins & Walker, 2000), classroom management skills (Woolfolk, Rosoff, & Hoy, 1990), pedagogical beliefs (Fives & Alexander, 2004), and reduced teacher stress (Greenwood, Olejnik, & Parkay, 1990). Soodak and Podell (1994) found that personal teaching

efficacy influenced teachers' tendency to take personal responsibility for teaching difficult students. Ashton and Webb (1986) found that teachers with high personal and general teaching efficacy assumed more responsibility for their students' learning, while Huang, Liu, and Shiomi (2007) found that teacher efficacy positively correlated to teacher self-esteem. Tschannen-Moran and Woolfolk Hoy (1998) and Ross (1994) have provided comprehensive reviews of teacher efficacy illustrating the importance of this construct in education.

Despite the popularity of teacher efficacy studies, controversy has surrounded the measurement of the construct because of the merging of the self-efficacy and locus of control theories. Several authors (Guskey & Passaro, 1994; Henson, Kogan, & Vacha-Haase, 2001; Tschannen-Moran & Woolfolk Hoy, 2001) have expressed concerns about the psychometric properties of the instruments used to evaluate teacher efficacy. Measurement uncertainties provide the historical context for the current investigation.

History of Teacher Efficacy Measurement

In a large survey evaluating numerous educational variables, a Rand Corporation Study (Armor et al., 1976) included the following two items: (a) "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." (b) "If I try really hard, I can get through to even the most difficult or unmotivated students." In the Rand investigation, teachers were asked to rate their level of agreement with the two questions (Armor et al., 1976). The results of the Rand study suggested a strong relation between the two-item teacher efficacy scale and teachers' success in teaching reading to urban, minority children. Utilizing the same questions within another Rand Corporation study, Berman

and McLaughlin (1977) found teacher efficacy to be predictive of the continuation of federally funded projects after the funding had ended.

The Rand researchers uncovered two conceptual strands of teacher efficacy, one measuring general teaching efficacy (Question 1) and the other personal teaching efficacy (Question 2; Armor et al., 1976). Teachers who strongly affirmed Rand question 1 indicated that factors outside themselves overwhelmed their power as teachers to influence the learning of students. This notion of general teaching efficacy (GTE) refers to belief in the power of teaching to overcome external factors, such as student socioeconomic status, value placed on education in the home, and family discipline (Tschannen-Moran et al., 1998). Teachers with high general teacher efficacy believe that good teaching can prevail over the influence of negative external factors that may impede student learning. Personal teacher efficacy (PTE) is defined as a teacher's belief in his or her personal teaching skills to overcome factors that could make learning difficult for a student (Tschannen-Moran & Woolfolk Hoy, 1998). Teachers who strongly agree with personal teacher efficacy statements indicate they have confidence in their own teaching skills to enhance student learning.

The Rand study items were based on the theoretical writings of Rotter (1966) and purported to relate to the internal/external locus of control dichotomy that is described within Rotter's social learning theory. According to this conceptualization of teacher efficacy, a teacher who has confidence in his or her ability to teach difficult students has an internal locus of control, whereas a teacher who believes that environmental factors (factors outside of themselves) have more influence on the reinforcement of teaching is thought to have an external locus of control. After the Rand study, researchers developed

longer, more reliable instruments within the Rotter tradition to tap the teacher efficacy construct (Guskey, 1981; Rose & Medway, 1981; Ashton, 1982). However, these measures were not widely adopted by other researchers interested in studying teacher efficacy (Tschannen-Moran & Woolfolk Hoy, 1998).

While scales were being developed in the Rotter tradition, Ashton and Webb began work with teacher efficacy and aligned the construct to Albert Bandura's social cognitive perspective of self-efficacy (Denzine, Cooney, & McKenzie, 2005). As stated earlier, self-efficacy describes a person's belief about his or her capability to achieve desired objectives (Bandura, 1977). Whereas the internal versus external locus of control perspective usually characterizes locus of control as a personality construct, Bandura believed that self-efficacy is domain specific and developed through four sources of information: mastery experiences, vicarious learning, social persuasion, and physiological arousal. Closely related to self-efficacy, Bandura also theorized about outcome expectations, which are an individual's expectation that he or she can successfully carry out a behavior to produce an outcome. In other words, outcome expectancy refers to an individual's estimation of the probable consequences of performing a task at his or her expected level of competence. Efficacy expectations refer to an individual's belief that he or she can complete the necessary actions to perform a given task (Bandura, 1986). Ashton and Webb postulated that general teaching efficacy corresponded to outcome expectations and personal teaching efficacy corresponded to efficacy expectations. Utilizing the social cognitive perspective, Ashton, Buhr, and Crocker (1984) created a measure employing a series of vignettes that described situations a teacher could encounter in the classroom. Like the aforementioned scales

created using the locus of control perspective, this measure never gained widespread acceptance in the teacher efficacy literature (Tschannen-Moran & Woolfolk Hoy, 2001).

In 1984, Gibson and Dembo created a longer teacher efficacy measure that would be more reliable and psychometrically sound than the original two-item Rand scale. Utilizing both the Rotter locus of control theory and the social cognitive conceptual framework, Gibson and Dembo, like Ashton and Webb, assumed that Rand Questions 1 and 2 coincided with Albert Bandura's (1986) outcome expectancy and self-efficacy dimensions, respectively. Their 30-item scale, the Teacher Efficacy Scale (TES), yielded a two-factor solution, which they suggested measured personal teaching efficacy (thought to correspond to the construct of self-efficacy), and general teaching efficacy (thought to correspond to the concept of outcome expectancy). Example questions from each of the respective scales are as follows: (a) If I try really hard, I can get through to even the most difficult or unmotivated students; (b) 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.

For almost two decades, the TES became the most popular teacher efficacy instrument within the educational literature (Ross, 1992; Woolfolk & Hoy, 1990). Utilizing the TES, many researchers (Emmer & Hickman, 1991; Soodak & Podell, 1996; Woolfolk & Hoy, 1990) found factor analytic results similar to those of Gibson and Dembo. But, in the mid-1990's, conceptual and statistical questions about the GTE and PTE constructs of the TES emerged (Brouwers & Tomic, 2003; Deemer & Minke, 1999; Emmer & Hickman, 1990; Guskey & Passaro, 1994; Kushner, 1993; Soodak & Podell, 1996; Tschannen-Moran & Hoy, 2001).

In 1994, Guskey and Passaro conducted a study utilizing the TES and concluded that the TES factors did not reflect personal and general teaching efficacy dimension, but rather an internal/external dichotomy of efficacy (“I can” versus “teachers can’t”). Deemer and Minke (1999), concerned with the positive and negative wording of the questions, investigated both a two-factor solution that reflected a positive and negative word orientation of the items and a four-factor fit that looked at the internal/external dimensions hypothesized by Guskey and Passaro (1994). Utilizing exploratory factor analysis, they found inadequate fit for both the two-factor and four-factor solutions. Kushner (1993) also failed to find the same factor structure as Gibson and Dembo when utilizing both exploratory and confirmatory factor analytic techniques.

Brouwers and Tomic (2003) investigated the TES using confirmatory factor analysis in which the four separate factor models previously suggested throughout its usage (Gibson and Dembo’s original two-factor model, Soodak and Podell’s and Woolfolk and Hoy’s three factor model, the three-factor model suggested by Emmer and Hickman, and a four-factor model that included the general teaching efficacy factor and the personal teaching efficacy factor divided into three subscales) were evaluated. They found that none of the models adequately fit the data and further suggested that the TES was “not suitable for obtaining precise and valid information about teacher efficacy beliefs” (p. 78). Denzine, Cooney, and McKenzie (2005) conducted confirmatory factor analysis of the TES and found that the two- and three-factor models that had been suggested by previous researchers did not adequately fit the data and that the adoption of the TES for so many studies of teacher efficacy had been premature. For a

comprehensive review of the early studies of teacher efficacy, refer to Tschannen-Moran & Woolfolk Hoy (1998).

Citing Bandura (1997), Tschannen-Moran, Woolfolk Hoy, and Hoy (1998) suggested that teacher's self-efficacy and locus of control are distinct constructs and their union may have contributed to the confusion in the measurement of teacher efficacy. Therefore, Tschannen-Moran and Woolfolk Hoy (2001) developed an instrument that would measure teachers' self-efficacy via a new theoretical model. Within their conceptual framework, teacher efficacy is defined as judgments about personal teaching and the teaching task. From this new theoretical perspective, which is most closely aligned with Bandura's social-cognitive perspective, Tschannen-Moran and Woolfolk Hoy developed the Teacher Sense of Efficacy Scale (TSES; formerly the Ohio State Efficacy Scale). Through a comprehensive instrument development process, the TSES measures the following: efficacy for instructional strategies, efficacy for classroom management, and efficacy for student engagement (Tschannen-Moran & Woolfolk Hoy, 2001). An example question from the TSES is as follows: *How much can you do to adjust lessons to the proper level for individual students?* Reliabilities (Cronbach's alpha) for the scales were reported as .94 for the total scale, .91 for efficacy for instructional strategies, .87 for efficacy for student engagement, and .90 for efficacy for classroom management. Construct validity was examined by correlating the TSES total score with Gibson and Dembo's Personal Teaching Efficacy Scale ($r = .64, p < .01$; Tschannen-Moran & Woolfolk Hoy (2001). Tschannen-Moran and Woolfolk Hoy (2001) indicate that researchers using the TSES should always conduct factor analysis utilizing their own data.

The three-factor structure of the TSES provides useful information regarding teachers' sense of self-efficacy, most closely related to the earlier conceptualization of personal teaching efficacy. However, the component of the construct closely linked to Rotter's locus of control conceptualization, is not directly measured. Therefore, no measure of teachers' beliefs in their abilities to overcome external obstacles to student learning is examined in this instrument. Some researchers suggest that this facet of teacher efficacy is difficult to measure reliably, may be a weaker predictor of actual behavior than self-efficacy (Bandura, 1997; Tschannen-Moran & Woolfolk Hoy, 1998), and is a theoretically separate construct than self-efficacy. Nonetheless, it remains an important component of teachers' beliefs about their abilities to bring about desired outcomes for their students. Many of the influential studies connecting teacher efficacy to important student and teacher outcomes were conducted utilizing measures that evaluated teacher efficacy through Rotter's theoretical framework (Armor et al., 1976; Ashton, 1985; Ashton & Webb, 1986; Berman & McLaughlin, 1977; Guskey, 1984; Smylie, 1988). Reason suggests that teachers who feel they can do little to increase the achievement of students, whom they believe have undesirable internal or external factors contributing to their academic difficulties, may interact differently with those students. Research further suggests that students are more engaged in learning and learn more when teachers take personal responsibility for their learning (Cooper & Tom, 1984; Firestone & Rosenblum, 1988). If teachers' beliefs about students' background characteristics are to be understood and modified through interventions, careful measurement of this construct should be undertaken. For this reason, the current researcher (in consultation with a University of Maryland research team) wrote the

Teacher Influence on Student Achievement scale (TISA). An example of a TISA scale question is as follows: *How much can you do to increase the achievement of a student from a disadvantaged family background?* The question structure of the TISA scale parallels the question structure of the TSES and asks teachers about their own individual beliefs, rather than teachers in general as with some previous scales attempting to measure the locus of control type of teacher efficacy.

Interventions that Affect Teacher Efficacy

Many studies provide evidence for the importance of teacher efficacy (despite measurement issues), but few researchers have studied interventions to alter teachers' sense of efficacy (Henson, 2002; Ross & Bruce, 2007). Ross (1994) suggested this problem stems from two sources: (a) the desire to link teacher efficacy to teacher characteristics that cannot be manipulated and (b) the supposition that teacher efficacy is an unalterable, global belief system. Because of the paucity of experimental teacher efficacy research, the results of the studies described above make causal arguments regarding efficacy difficult. It remains unclear what the direction of influence is regarding teacher efficacy in relation to the numerous variables purported to be related to the construct (e.g., student achievement). Is teacher efficacy malleable? Results from more general self-efficacy and locus of control literature are promising as they both appear to be alterable constructs (Betz & Schifano, 2000; Sofronoff & Farbotko, 2002; Hans, 2000; Hattie, Marsh, Neill & Richards, 1997). A few studies that attempted to alter teacher efficacy beliefs are described and critiqued below.

Standards-Based Mathematics Professional Development Program. In 2007, John Ross and Catherine Bruce conducted a randomized trial to determine the effects of a

professional development program in mathematics on teacher efficacy. Their study, conducted in a Canadian school district, included 106 sixth grade teachers who were randomly assigned to a treatment or control group. The professional development program consisted of a standards-based mathematics program in which treatment teachers received one full-day and three two-hour after-school sessions. The Teachers' Sense of Efficacy Scale (Tschannen-Moran & Woolfolk Hoy, 2001) was adapted for mathematics teaching and used as the efficacy measure for pretest and posttest. The reliability estimates for the TSES were high ($\alpha = 0.81 - 0.86$). Exploratory factor analysis suggested a good fit to Tschannen-Moran and Woolfolk Hoy's (2002) identified factors. Utilizing a multivariate analysis of covariance, the researchers found that pretests accounted for the most significant portion of variance in scores. While controlling for pretest, a treatment effect was found for the efficacy for classroom management scale. Although not reported in the manuscript, utilizing Cohen's *d*, the effect size was 0.58. No other significant effects were found for the other two efficacy scales, but the effect sizes for the efficacy for instructional strategies and engagement scales were 0.28 and 0.19, respectively.

Based on their results, Ross and Bruce (2007) concluded that teachers in the treatment group had more confidence in their ability to manage the mathematics classroom environment following the professional development. They hypothesized that because classroom management is often the prime concern for teachers (Fuller, 1969), it was logical that confidence in the use of new instructional techniques that engage students in mathematics led to increased confidence in classroom management abilities in mathematics. Although the Ross and Bruce (2007) investigation was encouraging, the

researchers utilized a homogenous sample and did not use an analytical technique that took into account the hierarchical nature of the data (i.e., teachers nested within schools).

Cognitive Coaching. Two studies that investigated the effect of Cognitive Coaching on teacher efficacy are described below. In the Cognitive Coaching program (Costa & Garmston, 1994), teachers are first paired with a coaching partner. Initially, the coach holds a planning conference to discuss the forthcoming lesson. The coach then helps the teacher define goals, strategies for teaching, procedures for data collection, and evidence of student achievement. Following an observation of the teacher conducting the lesson, the coach conducts a reflecting conference, shares the data with the teacher, and assists him or her in scrutinizing and applying it for future classroom teaching (Edwards, Green, Lyons, Rogers, & Swords, 1998). Communication skills during this progression are vital, with rapport building, questioning, paraphrasing, and probing being integral components of the process. When the progression is complete, the teacher and the coach exchange roles and the process begins again.

The goal of cognitive coaching is to increase teacher efficacy and create a more collaborative and professional teaching environment (Edwards et al., 1998). Operationally, Costa and Garmston (1994) defined cognitive coaching as a supervisor's utilization of strategies to alter a teacher's perceptions, decisions, and cognitive functions. The changes in thought processes will lead to changes in instructional behaviors, which will then lead to better student learning. The main theme of cognitive coaching, as described by Edwards et al. (1998), is that the thought processes of the teacher are the targets of change.

Edwards et al. (1998) investigated the effects of a Cognitive Coaching program, a Nonverbal Classroom Management program, and monthly dialogue groups on teacher efficacy while training to implement standards-based education over a three-year period. Their study utilized a non-equivalent quasi-experimental design with pretest and posttest. In addition to the Cognitive Coaching program described above, Edwards et al. also utilized Grinder's program of Nonverbal Classroom Management (Grinder, 1996). The Grinder program was developed by evaluating the nonverbal behaviors of teachers who were experts in managing classrooms. This program was created to improve teachers' nonverbal techniques to manage a classroom environment. The final component of this intervention study was the monthly dialogue groups in which teachers from various schools met to discuss the implementation of the new standards.

The participants included 440 kindergarten through twelfth grade teachers in a large school district in the western United States. Due to attrition, data were only available on 302 teachers at the end of the third year. The Teacher Efficacy Scale (TES; Gibson & Dembo, 1984) modified as suggested by Soodak and Podell (1996) to include an outcome efficacy subscale, was used to measure teacher efficacy. Factor analysis of this instrument was not conducted on the study data, but the reliability coefficients of the subscales were .82 (teaching efficacy), .81 (personal teacher efficacy), and .71 (outcome efficacy). The TES was administered before the initial training session, ten months after the initial training session, and 29 months after the initial training session. Repeated measures analyses of variance were used to determine differences between the treatment and control group and differences in the groups over time. Level of involvement in Cognitive Coaching and Grinder's Nonverbal Classroom Management program was also

evaluated by examining extent of use variables (i.e., frequency of paraphrasing). Results suggested treatment effects, with significant growth on the Teaching Efficacy subscale over time and between the treatment and control groups. However, while significant, the effects were small with eta-squared values of .03 (between groups) and .18 (within groups). The eta-squared value describes the proportion of variance in the dependent variable that is related to the independent variable. As with the Ross and Bruce study (above), Edwards et al. (1998) did not investigate the data utilizing multi-level modeling despite the hierarchical nature of teachers nested in schools. Additionally, lack of random assignment, high levels of attrition in the treatment group, and significant differences between groups on number of pre-service credits make the small effects found more problematic because each threatens internal validity.

Marrienne Krpan (1997) also investigated the influence of Cognitive Coaching on teachers' sense of efficacy. She used a nonequivalent comparison group design, with pretest and posttest. Teachers completed the Teacher Efficacy Scale (Gibson and Dembo, 1984) as modified by Soodak and Podell (1993). Although a comparison group was included in the study, the treatment and control group sizes were small ($n = 6$ and $n = 7$, respectively). Utilizing between group univariate analyses, no significant difference was found between the treatment and comparison group. An effect size calculated from data tabled in the report was of moderate size ($d = .55$) although its confidence band includes zero, so little information is in the results. Krpan also analyzed the within-group change by calculating a dependent samples t-test. The results were significant and the "effect size" was large ($d = 1.39$), but this difference cannot be interpreted as an effect.

Dare to be You Teacher Training. In 1995, Fritz, Miller-Heyl, Kreutzer, & MacPhee evaluated whether teachers' exposure to the Dare to be You (DTBY) prevention program training and curriculum would alter teachers' efficacy beliefs as measured by the Teacher Efficacy Scale (Gibson & Dembo, 1984). The DTBY program is a multi-level prevention program focused on the following: decision-making and problem-solving; assertive communication skills; responsibility and role modeling; and efficacy, esteem, and empathy for self and others. DTBY training puts emphasis on personal self-esteem, internal locus of control, and communication at several levels, including the teacher level. Teachers received 20-24 hours of experientially based training.

Fritz et al. (1995) utilized a nonequivalent comparison group with pretest design with 241 teachers matched on several variables, including age, gender, race, years of teaching, certification, level of education, professional memberships, educational magazine subscriptions, and prior training in drug and self-esteem curricula. The majority of teachers were at the elementary school level. The treatment and comparison groups were well-matched on all variables except that educational level and number of professional memberships were both higher for the comparison group. Because of the measured group differences, level of education was used as a covariate for the analyses. Teachers completed the TES at pretest, posttest, and a nine-month follow-up. The researchers also measured the level of use of the DTBY curriculum package for the treatment group.

Repeated measures multivariate analysis of variance indicated group by time interactions for personal and general teaching efficacy. Specifically, modest increases

were noted for the treatment group in terms of personal teaching efficacy, while large declines were reported for the comparison group. The strongest positive effects were noted for teachers who reported frequent usage of the DTBY program. General teacher efficacy scores significantly decreased for both the experimental and control groups, with a larger decrease reported for the comparison group. Effect sizes were not reported. Means and standard deviations were not included in the report; therefore it was difficult to estimate the size of the effect and to understand the practical implications of the intervention.

Responsive Classroom Approach. Rimm-Kaufman and Sawyer (2004) evaluated whether the implementation of the Responsive Classroom (RC) approach contributed to improvements in teachers' self-efficacy beliefs. RC is a teaching approach that combines social and academic learning throughout the day. The program stresses teacher empathy and the employment of classroom structure, with the main goal being the development of student self-control. Six integral components of RC are the (a) morning meeting, (b) rules and logical consequences, (c) academic choice, (d) guided discovery, (e) classroom organization, and (f) communication with parents. Program details are provided by Rimm-Kaufman and Sawyer (2004).

For this study, three schools were chosen to fully implement the RC process. These schools were chosen partially based on the teachers' interest in the RC program. Three comparison schools were also included. Two of the measures utilized with the sample of 69 kindergarten through third grade teachers included a classroom practices measure that evaluated the classroom teachers implementation of the RC approach and a self-efficacy measure adapted from Bandura (1993). The four subscales of the measure

were disciplinary self-efficacy, instructional self-efficacy, efficacy to create a positive school environment, and efficacy to influence decision-making. The overall scale had an alpha reliability of .91, with subscale reliabilities ranging from .65 to .89.

Utilizing step-wise regression analyses, the authors examined whether teachers who reported greater utilization of RC practices also reported higher self-efficacy. After controlling for teacher and school characteristics within the regression models, the researchers found that teachers reporting higher levels of RC practices also reported greater disciplinary self-efficacy, greater efficacy to create a positive school climate, and greater efficacy to influence decision-making. Although teachers who reported higher RC practices reported higher instructional self-efficacy, this result was not statistically significant. The findings show that small to large effects were detected. However, using a posttest only quasi-experimental design leaves the results open to numerous threats to internal validity. In fact, the researchers conclude their study suggesting that random assignment and/or pretests would have improved the study design, as the results they found could simply be spurious correlations.

The reviews of the four programs above (Standards-based Mathematics Professional Development Program, Cognitive Coaching, the Dare to be You Teacher Training program, and Responsive Classroom Approach) indicate that teachers' sense of self-efficacy can be altered through specific interventions. Each of the programs found some positive effect. Results from the above studies suggest that an explicit effort to enhance teachers' instructional skills is a necessary component to improving teacher efficacy. In addition, programs that include participant interaction may also help to facilitate changes in self-efficacy (Ross & Bruce, 2007).

Instructional Consultation Teams

One intervention, Instructional Consultation Teams (IC Teams), may be suited to produce positive changes in teachers' sense of self-efficacy. IC Teams is an attempt to reorganize the traditional early intervention teams in American public schools. In contrast to the majority of school-based team models that focus on assessment of the deficits within a student, IC Teams are intended to focus on evaluating the student-teacher-task interactions that are contributing to the successes and difficulties that students encounter. IC Teams were created on the foundation of Instructional Consultation (IC; Rosenfield, 1987). IC is a consultative approach driven by a structured problem-solving model comprising the following stages: contracting, problem-identification and analysis, intervention design and planning, intervention implementation, evaluation and closure (Rosenfield, 2008). As an ecological model, the focus of IC is not determining what is wrong with the child, but rather what factors in the instructional environment and the tasks themselves are keeping the student from meeting his or her goals. An underlying philosophy of IC is the need for an alteration of beliefs. Specifically, a fundamental assumption of IC Teams is that the belief system of a teacher may need to change in order to positively affect student performance. This key element provides justification for the current study as teacher efficacy is hypothesized to be a belief system that will be influenced by an IC Teams intervention.

IC utilizes consultee-centered consultation as a mode of professional development, wherein the specific focus of change is the teacher (consultee). Within IC, the consultant and consultee work together to specify and solve a problem (Rosenfield, 2008). The communication skills of the consultant are integral to successfully co-

constructing the problem and solution within the IC context and to help define the problem in behavioral terms. Although communication is a broad concept, IC operationalizes this term to include such specifics as clarifying, paraphrasing, and perception checking (Rosenfield, 1987, 2004).

While instructional consultation is the engine that runs the process, IC Teams are the vehicle that delivers the implementation of this comprehensive program. IC Teams is a manualized program that utilizes the following three components: process variables, delivery variables, and evaluation design variables. The process variables include the collaborative consultative process for problem-solving, including the working relationship between the case manager (consultant) and teacher (consultee), the use of instructional and behavioral assessment, and the stages of problem-solving. The delivery system variables include team functioning, the process for requesting assistance, the use of case management, and the structured documentation of cases and student progress. Finally, the evaluative component of the program includes evaluation of training, implementation, and outcomes.

The overall goals of IC Teams are to enhance, improve, and increase student and staff performance (Gravois, Rosenfield, & Gickling, 2002). Specific objectives include the development of a support network within a school, enhancing teachers' skills, development of school-wide norms of collaboration and problem-solving, and utilization of data for classroom and school decisions. An IC Team is composed of school personnel, including the IC Team Facilitator, a building administrator, general educators, special educators, school psychologist, school counselor, health provider, social worker, and others. The IC Team Facilitator receives more advanced training and coaching in the

IC process than the other members of the team. He or she facilitates professional development for staff members and supports the ongoing training and development of the IC Team. Although Instructional Consultation is most often delivered to individuals, Rosenfield, Silva, and Gravois (2008) hypothesize that all teachers within the school will benefit from the implementation of IC Teams. The overarching goal of the model is to create a group of teachers who talk and reflect about their instruction. Further, an environment of collaboration will be developed and a culture of problem-solving will become the norm.

A seven step progression through the IC Team service delivery model includes the following: (a) teacher completes a brief request for assistance, (b) team member assignment as case manager, (c) contract for professional collaboration, (d) assessment of student(s)' entry level skills, (e) baseline and goal establishment and documentation, (f) classroom strategies development, demonstration, and implementation, and (g) ongoing data collection to determine progress toward goals (Gravois, Rosenfield, & Gickling, 2002).

Prior research on IC Teams suggests links to student achievement, goal-attainment, and reduced special education referrals (Levinsohn, 2000; Gravois & Rosenfield, 2002, 2006). A qualitative study by Costas, Rosenfield, & Gravois, 2001 suggested that IC Teams helped professionals develop skills in data-based decision-making and improved instruction, altered beliefs about student problems, and changed school staff mood and motivation. Another qualitative study evaluating IC Teams suggested that teachers went through several conceptual changes after being a part of the process (Knotek, Rosenfield, Gravois, & Babinski, 2003). Teachers shifted their focus

from global to specific problems and became more adept at utilizing information from data collection to guide decision-making about students' instruction.

IC Teams Hypothesized Effect on Teacher Efficacy. Bandura (1982) indicated four ways in which self-efficacy beliefs can be altered: mastery experiences, vicarious experiences, social persuasions, and physiological arousal. The concept of mastery experiences suggests that success will raise self-efficacy, while failure will reduce it. I hypothesized that teachers who participate in IC Teams by having a consultation case will develop new competencies and knowledge and have success in teaching students whom they previously found academically and behaviorally challenging. This success will increase teachers' beliefs in their abilities to improve student performance. Because IC Teams is a school-based model in which not all teachers necessarily participate, vicarious experience will also be an important key to changing large numbers of teachers' self-efficacy beliefs, especially during the initial years of the intervention. Bandura (1982) posited that seeing others have success will increase the observer's sense of self-efficacy. Therefore, if teachers see their peers have success through IC Teams, their efficacy beliefs may also be enhanced. Social persuasion from IC Teams facilitators, principals, and other teachers will also play a role in altering efficacy beliefs. Within IC Teams, it is the collaborative culture, the IC Teams facilitator, and the instructional skill enhancements that will be integral components and mechanisms for changes in teacher efficacy. The final way in which efficacy beliefs can be altered is through physiological arousal. Physiological responses to situations can enhance or hinder efficacy beliefs. These responses generally play a smaller role in changing self-efficacy (Bandura, 1997)

and seem less relevant to the IC Teams Intervention. Appendix A illustrates the theoretical model of teacher change.

Gregoire (2003) suggested that belief change occurs when teachers see the value of a particular innovation and how it is relevant to their own lives. If this is true, IC Teams could alter teachers' efficacy beliefs because improving educational achievement for all students is relevant to teachers. Middleton (2002) hypothesized that the creation of a non-threatening environment where teachers can evaluate their identities, beliefs, and experiences will promote belief change. Within IC Teams, unlike many other school-based team models, teachers are not required to sit in front of a large group of colleagues and describe the difficulties they are encountering. Instead, the consultant meets with the teacher individually and then serves as a representative of the dyad within the team meetings. This model may be less threatening to teachers than many other traditional school team-based models and therefore will enable greater internal teacher changes.

Consultation and Teacher Efficacy

School-based consultation has become a prominent mode of service delivery in the school setting (Reschly, 2008; Tilly, 2008). Consultation is often conceptualized as a preventative approach and therefore differentiated from the traditional refer-test-place model of special education. Because of this, teachers who prefer consultation versus the traditional model of referral for standardized assessment may have differing personal and general teaching efficacy beliefs. Specifically, teachers who feel confident in their personal teaching may feel less need to consult with others. On the other hand, teachers with low general teaching efficacy may feel that consultation is unnecessary as external

factors are the cause of the student problem and therefore consultation would not be successful.

Research in this area is mixed. Fox (2004) demonstrated a significant positive relationship between general teacher efficacy and teachers' willingness to participate in consultation. This research also provided evidence for a negative relationship between general teacher efficacy and a teacher's likelihood of referring a student for special education via the traditional model. Personal teacher efficacy had no significant relationship to either variable. DeForest and Hughes (1992) found that teachers with high personal teacher efficacy rated consultants as more effective and interventions as more satisfactory when compared to teachers with low personal teacher efficacy. Although this research provides substantiation for a relationship between teacher efficacy and consultation, teacher efficacy was the antecedent to participating in consultation. For the current study, participation in a school-wide consultation package is the hypothesized precursor to changes in scores on efficacy instruments.

Three studies evaluated the influence of consultation on teacher efficacy. Deaver (2005) found that collaborative consultation positively affects teacher efficacy. She utilized a qualitative approach in which she coded transcripts from interviews. From the transcriptions, she concluded that improvements were made in teachers' efficacy beliefs. The second study, conducted by Alkon, Ramler, and MacLennan (2003) was an experimental study conducted in urban childcare settings. As part of this study, the researchers wanted to determine whether duration of mental health consultation was associated with teacher efficacy. The Teacher Opinion Survey (Geller & Lynch, 1999), a 13-item scale designed to evaluate teachers' confidence in working with children with

challenging behaviors, was the teacher efficacy measure. The reliability of the scale was moderate ($\alpha = .66$). Results of this evaluation suggest that mental health consultation implemented for over a year produced a significant increase in teacher efficacy. Qualitative results from the study indicated a shift from external to internal locus of control. Following the mental health consultation, teachers felt more prepared to help children with emotional and behavioral problems that may be due to family or environmental factors (Alkon, Ramler, & MacLennan, 2003).

Vu et al. (2009) assessed the effect of IC Teams on several teacher outcomes, including teacher efficacy. Using the same data as the current study, Vu et al. (2009) found that the intervention significantly increased teachers' sense of efficacy (coefficient = .10, $p < .05$). For their study, however, the researchers evaluated teacher efficacy, in addition to several other teacher variables, as a unidimensional construct within a two-level HLM model. They did not separately test the effect of the intervention on the TISA items that were written specifically for this study. The current evaluation differs from the Vu et al. (2009) study by examining the two teacher efficacy scales in a three-level model that considers the measurement dichotomy of the teacher efficacy construct (described below). Additionally, for the current study, teacher demographic variables were utilized as covariates to potentially increase power. The multivariate method was also used as a way to potentially increase power and to evaluate any differential effect of IC Teams on the separate scales.

Research Questions and Framework

A study of the effectiveness of an intervention to increase teacher efficacy must employ a research design that will allow causal inferences and employ useful

measurement instruments. For this reason, the current study is designed to assess whether the implementation of IC Teams increases teacher efficacy through two conceptual lenses. Specifically, the current investigation will evaluate the effect of IC Teams on teacher efficacy via a true experimental design and two alternative efficacy-related measures.

This research is designed to answer the following specific questions:

1. Is there evidence that the TSES and TISA scales measure distinct dimensions of the teacher efficacy construct?

2. How effective are Instructional Consultation Teams at increasing teacher efficacy as measured by the TSES Efficacy for Instructional Strategies scale and the TISA scale?

3. Are teacher demographic variables predictive of teacher efficacy as measured by the TSES and TISA?

METHOD

This research is part of a larger longitudinal randomized control trial evaluating the effects of IC Teams on several school, teacher, and student outcomes. In 2005, 34 ethnically and linguistically diverse schools agreed to be part of the large IC Teams study and were then randomly assigned to treatment ($n = 17$) or control ($n = 17$) groups. Baseline data were collected during the 2005-2006 academic year. For the current study, outcome data were used from the 2007-2008 academic year, which is the second year of

IC Teams implementation. The present investigation is focused on the effect of IC Teams on teacher efficacy.

Subjects

Subjects are 1203 teachers (kindergarten through fifth grade) in 34 elementary schools within a large suburban school district. The majority of teachers are White, non-Hispanic women with teaching experience ranging from less than one year to more than 20 years. The researchers used the following definition to help guide the selection of participants, “anyone who teaches more than one student.” Responses from regular education, special education, and specialist teachers were included.

Data Collection Procedures

A teacher survey was administered to participants via an online format. For each year of the data collection (2006, 2007, and 2008), the same online survey format was used. The overall survey consisted of 94 questions that asked teachers about their efficacy, instructional practices, collaboration with others, job satisfaction, and general demographics. Teachers were also asked about the organizational focus of their schools. The completion of the survey was voluntary and participants were required to consent prior to taking the survey. All data were collected so that the researchers had no way of linking the survey responses to individual respondents. Teachers were notified that information would be presented in an aggregated form. Each teacher asked to complete a survey was sent a hard-copy letter and a small token (notepad) prior to completing the survey. Three thank you/reminder emails were sent to teachers following the initial request in an effort to increase response rate. A focus group of five teachers was utilized

prior to the initial baseline survey. Feedback from the focus group suggested the entire survey took 10 to 20 minutes to complete. The response rate for the survey was 89%.

Measures

Teacher Efficacy Scales. Based on the initial (2005-2006 school year) data collection and decisions of key members of the program evaluation staff for the larger randomized experiment, the short-form Efficacy for Instructional Strategies scale (Appendix B; Tschannen-Moran & Woolfolk Hoy, 2001) was retained as the TSES measure for the longitudinal study. The Efficacy for Instructional Strategies short form, which includes six questions, was adapted from its original form by changing the response categories from nine to five options and by including two additional questions from the TSES-long version. The five response categories of the modified TSES were consistent with the original response categories of the measure (with the exception of “not at all” being added to the first anchor response): (1) nothing/not at all; (2) very little; (3) some; (4) quite a bit; and (5) a great deal. The additional two questions from the TSES long form were used because of the possible relevance of these questions to the experimental evaluation of the IC Teams intervention package. In a previous study, the reliability estimate (coefficient alpha) of the TSES Efficacy for Instructional Strategies sub-scale was .86 in a sample of 410 pre-service and in-service teachers, all of whom were students at three major universities or teacher volunteers from elementary, middle, and high schools (Tschannen-Moran & Woolfolk Hoy, 2001).

The TISA scale (Appendix C) was originally prepared during the baseline data collection after reviewing previous measures that utilized the concept of general teaching efficacy and locus of control (Armor, et al., 1976; Gibson & Dembo, 1984) and through

contemplating teacher outcome expectations of the larger study of IC Teams. These questions ask teachers how much they can do to improve the achievement of students in the face of traditionally negative external factors. The response categories for the TISA scale paralleled the response categories for the TSES. The reliability estimate (coefficient alpha) from the baseline data was .85. During the 2006-2007 data collection, four new questions were added to the original TISA to enhance its measurement utility for the present study. In conjunction with other University of Maryland researchers, I wrote the four new questions based on theoretical hypotheses regarding teachers' beliefs about increasing achievement of students from specific backgrounds. The reliability coefficients (coefficient alphas) for the TSES Efficacy for Instructional Strategies subscale and the TISA scale for the current data collection were .85 and .92, respectively.

Exploratory factor analysis was conducted to determine the underlying factor structure of the TSES Efficacy for Instructional Strategies scale and the TISA scale for the current data set. The two measures were analyzed together to determine whether separate constructs were evident. Based on recommendations from Russell (2002) and Costello and Osborne (2005), principal axis factoring with varimax rotation was utilized. This procedure was also employed for the previous studies in which the TSES was developed (Tschannen-Moran & Woolfolk Hoy, 2001) and therefore provided consistency in analysis.

Specifying too many or too few factors is problematic (Zwick & Velicer, 1986). Although there is no consensus on the criteria to use for determining the number of factors to retain (Hayton, Allen, & Scarpello, 2004), evidence suggests that parallel analysis (Horn, 1965) may be the most defensible method for this determination, though

it is not commonly seen in behavioral research studies employing factor analysis (Russell, 2002; Zwick and Velicer, 1986). Parallel analysis is a technique whereby eigenvalues from a set of data (prior to rotation), which measure the number of units of variance explained by a particular factor, are compared to a matrix of random values with the same number of variables and cases (Franklin, Gibson, Robertson, Pohlman, & Fralish, 1995). Eigenvalues from the actual data set that are larger than the eigenvalues from the corresponding random data are retained. Because many researchers employ Kaiser's criterion, which suggests retaining factors with eigenvalues over one, the use of Catell's scree plot, which suggests plotting the eigenvalues to determine where they level off or become scree, or a priori hypotheses regarding the number of factors to retain, multiple methodologies were used to determine the number of factors to retain on both scales for this study. The use of multiple procedures to determine the number of factors to retain is desirable, because no decision rule is completely accurate (Thompson & Daniel, 1996).

When evaluating the scree test, Kaiser's eigenvalue criteria, and parallel analysis, principal axis factoring yielded a two-factor solution that accounted for 52.2% of the variance. Prior to rotation, the first factor accounted for 47.83% of the variance, while the second factor accounted for 10.22% of the variance. Following rotation, the first factor accounted for 30.9% of the total variance, while the second factor accounted for 21.3% of the variance. Within the parallel analysis, two eigenvalues from the data set were larger than the corresponding eigenvalues from the randomly generated data set. The two-factor solution corresponded to separate TSES and TISA scales. A criterion of .5 or higher was utilized to determine if an item should be retained in a factor-based scale. A

factor loading denotes the correlation between the variable and the factor. Therefore, utilizing a .5 minimum criterion for a factor loading means that at least 25% of the variance is accounted for by the factor. For the two theoretically-based scales, all factor loadings were above .5. Although one question showed a high cross-loading (.42), it was retained on the factor to which it theoretically corresponds. Communalities (h^2), which measure a variable's shared variance with other variables, are reported in Table 1. All were moderate in size.

Teacher Demographics. As part of the online survey, teachers were asked questions about their years of experience, professional role, educational background, ethnicity, and gender.

Intervention Design and Implementation

The IC Teams intervention is a comprehensive package implemented within schools to enhance improve, and increase student and staff performance. Explicit procedures and guidelines for the content, process, training, implementation, and evaluation of this program are included in a training manual and published text (Gravois, Rosenfield, & Gickling, 2002; Rosenfield & Gravois, 1996). The overall longitudinal experiment is in its third year and the IC Teams program is in its second full-year of implementation in the 17 treatment schools. The training and implementation process for IC Teams are based on Fullan's (1991) three phases of school change: initiation, implementation, and institutionalization.

During the first half of year one (Fall 2005), treatment schools were in Phase 1 or the initiation phase (Fullan, 1991), which entailed the collection of baseline data and readiness activities provided to the treatment schools. These activities included

identification of IC Team facilitators, advanced training for the facilitators, principals, and another key school staff member (buddy) in each of the treatment schools. The IC Teams facilitator also received seven additional two-day training sessions and feedback on a coached case. Mid-way through the first year, the IC Teams were formed in each treatment school and a 3-day training, led by the school IC Teams facilitator, was held for team members. This began Phase 2, which is the implementation phase. The facilitator and team members then practiced collaborative communication skills, systematic problem-solving, instructional assessment, and behavioral assessment with volunteer teachers.

During the second year, the IC Teams delivery system was more fully implemented within the treatment schools. Teachers in the treatment schools began consulting with case managers on the team concerning real student concerns and teams met weekly for training led by each school's IC Team facilitator. The Laboratory for Instructional Consultation Teams at the University of Maryland provided technical support. During the current year of the project (the third year), the teams continue to implement the intervention as they did in the previous year. Schools will be considered to be in Phase 3 (institutionalization) when they implement the process with high levels of fidelity for two or more years according to criteria for institutionalization provided by program implementers.

Data on quality of implementation for the 2007-2008 academic year were collected using the Level of Implementation (LOI) procedure (Fudell, 1992; Gravois, Fudell, & Rosenfield, 2005; McKenna, 2005). The LOI scale combines interviews, record reviews, and observations of selected IC Team case managers and team members.

The LOI data indicate that personnel in treatment schools were capable of high quality program implementation. Scores on the LOI scale for sampled cases ranged from 88% to 100%. Two schools did not have LOI or utilization scores (described below) as they had large amounts of missing data. Utilization data, which refers to how many teachers in the treatment school ever consulted with the IC Team, indicates that the percentage of teachers accessing the IC Team in treatment schools ranged from 34% to 93% with an average of 65% of teachers (Berger, Vaganek, Yiu, Hong, & Nelson, 2009). The utilization data refer only to regular classroom teachers. Special educators and other specialists were not included in this data collection. Table 2 shows the implementation quality and level of use data for the 17 treatment schools.

Research Design

The study design is a randomized experiment. Specifically, a pretest-posttest control group design, with schools as the unit of random assignment, was used. Random assignment eliminates, in expectation, the selection threat to internal validity and the probability of other threats to internal validity is reduced. The fundamental strength of using random assignment is that confounding variables are unlikely to be correlated with treatment conditions (Shadish, Cook, & Campbell, 2002). Vu et al. (2009) conducted analyses on baseline data for the overall study and determined that no significant differences existed between treatment and control schools in terms of teacher and school demographics on the survey scales.

Power Analysis

Before the randomized experiment began, Rosenfield and Gottfredson (2004) estimated that power would equal .80 to detect treatment effect sizes of .22 for outcomes.

Due to the design of the intervention, schools are the unit of assignment to treatment. Because of this, the number of schools in the study (34) is the most important limiting factor on statistical power.

Multilevel Analysis

Because schools and not teachers are the unit of assignment and treatment in this study, the school was the unit of analysis for treatment effects and hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) was utilized. The use of HLM helps to avoid underestimation of standard errors, aggregation bias, and heterogeneity of regression problems. Underestimation of standard errors occurs when individuals within particular schools tend to have similar characteristics. HLM corrects this problem by incorporating random error for the organizational units. Aggregation bias is the group-induced or aggregation-induced correlation among variables that misrepresents or misestimates the individual-level correlation. HLM helps to eliminate this bias by parsing relationships between variables into separate level components. Finally, heterogeneity of regression occurs when the relations between individual characteristics and outcomes differ across schools. With HLM, researchers can estimate individual regression coefficients for each unit of an organization and model variation among the organizations (Raudenbush & Bryk, 2002).

Analytic Approach

Descriptive analyses were conducted to evaluate the average scores on each of the outcome variables by treatment condition. The results of this analysis (shown in Table 3) indicate that the teachers in the treatment sample scored slightly higher than those in the control sample on both the TISA and TSES Efficacy for Instructional Strategies scales.

Outcome variables were screened for outliers. For the TSES Efficacy for Instructional Strategies scale, eight outliers (six cases from treatment schools) were identified to be more than three standard deviations below the mean, while the TISA scale had three outliers (two cases from treatment schools) by the same definition. Two of the eleven cases were outliers on both scales, so nine cases in total were affected. These nine aberrant cases were evaluated and deleted from the file. This decision was made to ensure that the outcome variables would more closely approximate normal distributions and thus meet a fundamental assumption of the statistical analyses. Analyses were also conducted with the outliers included to evaluate the effect of their deletion.

Several teacher-level variables that were collected as part of the overall survey or through school records were evaluated. Polytomous categorical variables were recoded into meaningful dummy variables based on descriptive and means analyses. Table 4 provides information regarding the teacher-level predictors and their coding.

Zero-order correlation matrices were produced to evaluate the relations between the two teacher efficacy scales and the available teacher demographic variables. Table 5 provides information about the correlations. For the TSES Efficacy for Instructional Strategies scale, a pretest of teacher efficacy questions (taken during baseline data collection), age of teacher, educational level, years as a teacher, and years at a particular school were significantly correlated to the outcome. The pretest, gender, job type, educational level, and years as a teacher variables were significantly correlated with the TISA scale. Variables that were not significantly correlated with the outcomes were dropped from further analyses and were not included in the subsequent HLM models.

Multicollinearity of the predictor variables was explored. All variance inflation factors (*VIF*) were below seven and partial correlations were equal to or below zero-order correlations.

To aid in interpretation, the outcome variables and the age in years variable were transformed into *z*-scores by subtracting the grand mean from the score and dividing it by the standard deviation. All other predictor variables were dichotomous. The distributions of the outcome variables were evaluated for normality. The TSES Efficacy for Instructional Strategies scale had a significant negative skew as determined by calculating a skewness *z* statistic from the information provided by the SPSS output. This skewness *z* statistic was calculated by dividing the skewness statistic (-.342) by the skewness standard error (.07). This value (4.82) was higher than the suggested acceptable value of 3.3 (Tabachnick and Fidell, 2001). Because of the significant skew, several transformations were attempted. Computation of the square root of the reflected scale produced the distribution most closely approximating the normal and thus was utilized. The TISA scale had a skewness *z* value below 3.3, so no transformation was deemed necessary.

An *intent to treat* approach was used for the analyses. Data from the baseline year were evaluated to determine the teachers intended to receive the treatment or to serve as controls. In all, 1212 teachers were working in the 34 experimental schools at the time of school random assignment, and so are included in the intent to treat group. Data from the current year data (2007-2008) were merged with the baseline data. Of the original intent to treat group, 41% (496 teachers) had no data for the 2007-2008 school year. These missing values were imputed using SPSS Missing Value Analysis. This

program utilizes the EM algorithm to impute values, so an assumption was made that the data are missing at random (Little & Rubin, 2002). To evaluate the effect of the imputations, the data were also analyzed excluding those cases with missing data.

Factor analysis was conducted on the data following the imputation. Results were consistent with the initial factor analysis in which two separate factors were evident. This provided continued support for the need to analyze the two scales as separate entities. In addition, the correlation between the two scales was calculated. The total score on the TISA scale was positively related to the TSES Efficacy for Instructional Strategies scale score ($r = .66, p < .01$). Based on calculated alpha levels for each scale, the estimated true score correlation is .75, indicating that 44% of the variance in the true scores does not overlap.

Attrition and Treatment Diffusion

Analyses were completed to determine if the treatment and control groups differed in terms of attrition. This analysis was conducted by utilizing treatment status and teacher variables (i.e., age, gender, and ethnicity) to predict missingness. Specifically, a dichotomous outcome model in HLM was constructed to determine if treatment condition predicted attrition. The dichotomous dependent variable was a one or zero depending on whether data were present or missing. Level 1 predictors were also used to see if teacher characteristics predicted attrition. Treatment diffusion, defined as teachers moving from one experimental condition to the other, was also evaluated.

Multilevel, Multivariate Model

Because exploratory factor analyses indicated there are two teacher efficacy scales (TSES Efficacy for Instructional Strategies scale and the TISA scale), a three-

level model was used to incorporate the two teacher efficacy measures. This model is referred to as a multivariate, multilevel model (Sniders & Bosker, 1999; Hox, 2002). At level one, both latent constructs were simultaneously estimated. Although it is possible to evaluate the two dependent variables separately, Sniders and Bosker (1999) suggest the following reasons for joint analyses: (a) Information can be obtained about the correlations between the dependent measures and how much the correlations depend upon the teacher and school level. (b) Because the dependent variables are highly correlated in univariate analysis, the use of a multivariate design may produce smaller standard errors and potentially increase power. (c) A multivariate analysis is needed to evaluate whether the effect of a predictor variable is larger on one of the dependent variables when the data were taken from the same individuals.

The equations below detail the fully unconditional three-level model. The unconditional model has no predictors included at any level. There is no intercept or error term at the first level as this level is constructed only to define the multivariate structure of the model (Hox, 2002). The notation follows that used by Raudenbush and Bryk (2002). The unconditional model equations (1,2,3,4,5) are as follows:

The Level 1 (measures within teachers) is:

$$Y_{ijk} = \pi_{1,jk}a_{1,jk} + \pi_{2,jk}a_{2,jk}$$

(1)

where

Y_{ijk} is the response on scale i of teacher j in group k ;

$\pi_{1,jk}$ is the score for teacher j in school k on scale 1;

π_{2jk} is the score for teacher j in school k on scale 2;

$i = 1, 2$ for each of the teacher efficacy scales included in the model;

a_{1jk} is an indicator variable that takes a value of 1 for scale 1 and 0 otherwise;

a_{2jk} is an indicator variable that takes a value of 1 for scale 2 and 0 otherwise;

The Level 2 (teachers within schools) model is:

$$\pi_{1jk} = \beta_{10k} + r_{1jk}$$

(2)

$$\pi_{2jk} = \beta_{20k} + r_{2jk}$$

(3)

where

β_{10k} is the mean score on teacher efficacy scale 1 in school k ;

β_{20k} is the mean score on teacher efficacy scale 2 in school k ;

r_{1jk} is the random effect of scale 1 associated with teacher jk .

r_{2jk} is the random effect of scale 2 associated with teacher jk .

The Level 3 (between school) model is:

$$\beta_{10k} = \gamma_{10k} + u_{10k}$$

(4)

$$\beta_{20k} = \gamma_{20k} + u_{20k}$$

(5)

where

γ_{10k} is the grand mean score of scale 1;

γ_{20k} is the grand mean score of scale 2;

u_{10k} is the random error at the school level for scale 1;

u_{20k} is the random error at the school level for scale 2.

The unconditional three-level model provides information regarding the reliabilities of each scale, the amount of between school variance in each scale, and the correlations of the scales at both the teacher and school level.

Discriminant validity was evaluated by examining the correlations between the constructs while disattenuating for measurement error at the teacher and school levels. The correlations between the teacher efficacy latent variables at the teacher and school level were expected to be higher than the correlation between the manifest scale scores. These correlations indicate the extent to which the constructs tapped by the TSES and TISA scales overlap. The procedures described above were adapted from Raudenbush, Rowan, and Kang (1991), Miller and Murdock (2007), Tate and Pituch (2007), and Hox (2002). To use the above approach for the analyses, the Level-1 error variance was set to approximate zero (.0001) and the Level-1 intercept was deleted from the model as suggested by Tate and Pituch (2007) and Hox (2002) for multivariate, multilevel models.

A level two, teachers within-schools model, was then created by adding covariates at

level 2 while leaving level 3 fully unconditional. The covariates used were those variables that were significantly correlated to each outcome variable within an exploratory ordinary least squares analysis. All intercepts and slopes were tested to decide whether to use fixed or random effects as suggested by Raudenbush and Bryk (2002). Initially, the predictors were group-mean ($X_{ij} - X_{.j}$) centered and specified to have random slopes to determine if the relationship between the dependent variables and the predictor variables differed among schools. If the relationships did not vary

significantly, the variables were grand-mean ($X_{ij} - X_{..}$) centered and slopes were fixed to zero for subsequent modeling. Finally, the significance of each predictor variable was tested and evaluated for its usefulness in the model. Only those variables that were significant predictors of the outcomes were retained in the model.

The Level 3 model is the between-schools model. The level-two intercept and slope coefficients were the dependent variables in the level-three equations. Intervention status was added to the Level 3 uncentered to determine if it was a significant predictor of the outcome variables. Due to the complex nature of the three-level, multivariate model, no other covariates were added at the school level. This decision was also made because the main questions for this study pertain to the effect of the treatment on the outcome variables, and treatment is uncorrelated in expectation with all other school-level variables by virtue of random assignment.

Finally, analysis was conducted to evaluate the effect of treatment on both outcomes simultaneously using general linear hypothesis testing as suggested by Tate and Pituch (2007). The hypothesis test evaluates whether the treatment effect for *both* outcomes is statistically significantly different from the null hypothesis. In other words, it is a simultaneous test that the two fixed effects are both equal to zero.

RESULTS

Attrition and Treatment Diffusion

Although considerable attrition occurred from baseline to the 2007-2008 academic year (41% of teachers from the baseline year did not have values for the current year data), treatment condition was not a significant predictor of attrition ($p = .18$). Furthermore, gender, age of teacher, ethnicity, and pretest did not predict attrition. There

were no significant interactions between treatment and the Level-1 predictor variables. Little mobility between conditions was found. Thirteen teachers moved from treatment to control and seventeen teachers moved from control to treatment between the baseline and current year data collection. These are small percentages of teachers given the large sample sizes.

Unconditional Model

Results from the unconditional model indicated intra-class correlation coefficients (ICC) of .02 and .03, for the TSES Efficacy for Instructional Strategies and TISA scales, respectively. This means that approximately 2% of the variation for the TSES Efficacy for Instructional Strategies scale lies between schools, while 3% of the variation for the TISA scale lies between schools. The school-level reliability coefficients reported from HLM6 are a function of the between school variance and the number of teachers within each school for the items. For the TISA scale the average school-level reliability was moderate (.54), while the average TSES Efficacy for Instructional Strategies reliability was low (.36). These values are moderate to low because the between school variance and the number of teachers within each school were relatively small.

Correlations at Level-2 and Level-3

Correlations between the two scales were evaluated through the unconditional three-level model. Results are presented in Table 6. As expected the correlation between the scales when correcting for school-level unreliability was larger than the uncorrected correlation. The teacher-level correlation was similar whether or not evaluated in a multi-level model because the intraclass correlation was small.

Within-School Model

Initially, at Level-2, the pretest variable was specified as grand-mean centered and fixed to serve as a control variable for each outcome. All other teacher variables that were significant predictors of the two teacher efficacy outcomes when examined using ordinary least squares regression (gender, job type, educational level, years of experience, age in years, and years at school) were group-mean centered and specified to have random effects to determine whether these variables were significantly related to the outcome variables and whether the relationships significantly varied between schools within HLM6. Results indicated that for the TISA scale, pretest (coefficient = .59, $SE = .02$, $p < .01$), and gender (coefficient = .31, $SE = .07$, $p < .01$) were significantly correlated to the outcome. The relationship between gender and the TISA scale did not vary significantly between schools, so slopes were fixed. According to the results, male teachers, on average, scored .31 standard deviation units higher than female teachers on the TISA scale. For the TSES Efficacy for Instructional Strategies scale, pretest (coefficient = -.56, $SE = .03$, $p < .01$) and years at school (coefficient = -.14, $SE = .04$, $p < .01$) were significantly correlated to the outcome. Note that the TSES Efficacy for Instructional Strategies scale was transformed and therefore, coefficients with a negative sign should be interpreted in a positive direction. The relationship between the years at school variable and the TSES Efficacy for Instructional Strategies scale did not vary significantly between schools, so slopes were fixed. According to the results, teachers who have been teaching at a particular school six or more years, scored, on average, .14 standard deviation units higher than teachers who were at a school for less than six years. Results of the final within-schools model are displayed in Table 7.

Between-School Model

For the TISA scale, treatment status was entered uncentered as a predictor variable at the school level (Level-3), with the pretest and gender variables taken into account at Level-2. For the TSES Efficacy for Instructional Strategies scale, treatment status was also added uncentered at Level-3, with the pretest and years at school variables accounted for at Level-2. Treatment status significantly predicted the TISA scale (coefficient = .20, $SE = .06$, $p < .01$) and the TSES Efficacy for Instructional Strategies scale (coefficient = -.16, $SE = .05$, $p < .01$). The results indicate that teachers in the treatment condition scored .20 standard deviation units higher on the TISA scale and .16 standard deviation units higher on the TSES Efficacy for Instructional Strategies scale as compared to control group teachers, when accounting for Level-2 covariates. Effect sizes for the TISA and TSES Efficacy for Instructional Strategies scales were small (TISA $d = .22$, TSES Efficacy for Instructional Strategies $d = .16$). Results from the analyses that included outliers were consistent with the above results. Therefore, only results with the outliers removed are reported. The final between-school model can be conceptualized as reported below in equations 6 to 14. Table 8 summarizes the results of the final between schools model.

The Level 1 (scales within teachers) is:

$$Y_{ijk} = \pi_{1,jk}a_{1,jk} + \pi_{2,jk}a_{2,jk}$$

(6)

where

Y_{ijk} is the response on scale i of teacher j in group k ;

$\pi_{1,jk}$ is the score for teacher j in school k on scale 1;

π_{2jk} is the score for teacher j in school k on scale 2;

$i = 1,2$ for each of the teacher efficacy scales included in the model;

a_{1jk} is an indicator variable that takes a value of 1 for scale 1 and 0 otherwise;

a_{2jk} is an indicator variable that takes a value of 1 for scale 2 and 0 otherwise;

The Level 2 (teachers within schools) model is:

$$(7) \quad \pi_{1jk} = \beta_{10k} + \beta_{11k} (\text{Pretest}_{jk} - \overline{\text{Pretest}_{jk}}) + \beta_{12k} (\text{Gender}_{jk} - \overline{\text{Gender}_{jk}}) + r_{1jk}$$
$$(8) \quad \pi_{2jk} = \beta_{20k} + \beta_{21k} (\text{Pretest}_{jk} - \overline{\text{Pretest}_{jk}}) + \beta_{22k} (\text{Years at School}_{jk} - \overline{\text{Years at School}_{jk}}) + r_{2jk}$$

where

β_{10k} is the mean score on teacher efficacy scale 1 in school k ;

β_{20k} is the mean score on teacher efficacy scale 2 in school k ;

β_{11k} is the effect of the Pretest on scale 1;

β_{12k} is the effect of the Gender on scale 1;

β_{21k} is the effect of the Pretest on scale 2;

β_{22k} is the effect of the Years at School on scale 2;

r_{1jk} is the random effect of scale 1 associated with teacher jk ;

r_{2jk} is the random effect of scale 2 associated with teacher jk .

The Level 3 (between school) model is:

$$(9) \quad \beta_{10k} = \gamma_{100} + \gamma_{101} (\text{Treatment}_k) + u_{10k}$$

$$\beta_{11k} = \gamma_{110}$$

(10)

$$\beta_{12k} = \gamma_{120}$$

(11)

$$\beta_{20k} = \gamma_{200} + \gamma_{201} (\text{Treatment}_k) + u_{20k}$$

(12)

$$\beta_{21k} = \gamma_{210}$$

(13)

$$\beta_{22k} = \gamma_{220}$$

(14)

where

γ_{100} is the grand mean score of scale 1;

γ_{200} is the grand mean score of scale 2;

γ_{101} is the effect of Treatment on scale 1;

γ_{201} is the effect of Treatment on scale 2;

u_{10k} is the random error at the school level for scale 1;

u_{20k} is the random error at the school level for scale 2;

Treatment takes on values of 0 if control and 1 if treatment.

Proportion of Variance Explained for the TISA and TSES Efficacy for Instructional Strategies

To evaluate the amount of variance the final model explains, the percent reduction in residual variances between the unconditional and final models was calculated. This information is obtained by subtracting the variance components in the unconditional

(null) model from the variance components in the final (full) model and dividing these by the variance components in the null model (Hox, 2002). When using these calculations, the final model explains 37% of variance within schools and 68% of variance between schools for the TISA scale and 32% of variance within schools and 93% of variance between schools for the TSES Efficacy for Instructional Strategies scale. See Table 9.

Multivariate Testing of Total Treatment Effect

Hypothesis testing was used within HLM6 to evaluate the multivariate test of the total effect of treatment on both outcomes. The null hypothesis for this test is

$H_0 :$

$$\gamma_{101} = \gamma_{201} = 0. \tag{15}$$

The results of the multivariate test were significant ($\chi^2 = 13.21, df = 1, p < .01$).

These results can be interpreted similarly to a MANOVA, indicating that the effect of IC Teams was significant for both outcomes tested simultaneously.

Analyses with Cases with Complete Data

In the model that excluded cases with missing values ($n = 695$), treatment was a significant predictor for both outcomes and the pretest covariate was significant at Level-2 for each outcome. The coefficient value for the TISA scale (.19) was similar to the result from the model with imputed values, but the coefficient value for the TSES Efficacy for Instructional Strategies scale was smaller (-.03). Specific results of this model are shown in Table 10.

To evaluate the differences in the coefficient values for the data with and without imputation, follow-up analyses were conducted. Descriptive analyses were used to evaluate the average scores on each of the outcome variables by treatment condition. The

results are shown in Table 11. The mean scores on each scale by treatment condition were similar to those reported for the imputed data (Table 3). However, the standard deviations were higher for the non-imputed data, which may account for the differences in the coefficient values in the imputed and non-imputed data sets. The values imputed via the use of the EM algorithm used are optimal (predicted) values but they lack variability presumably present in the complete data (if they were available). That is, the imputed values lack a random error component (Enders, 2001). Schafer and Graham (2002) show that when data are not missing completely at random (MCAR) that complete data analysis (i.e., listwise deletion of cases missing some data) can produce biased estimates of variances and regression coefficients. However that may be, the analyses of these data using complete case analysis and using imputation were intended to provide a sensitivity test of the analytic assumptions. With respect to statistical significance of the effect, the main result survives the sensitivity test. With respect to the size of the effect however, the alternative assumptions produce different estimates of the effect size for the TSES EIS scale.

DISCUSSION

Treatment Effect

The purpose of this study was to assess the effect of IC Teams on teachers' sense of efficacy through two conceptual lenses. The main effects for IC Teams on both outcome variables were statistically significant. These findings indicate that a comprehensive teacher-focused, consultee-centered consultation package (IC Teams) significantly increased teachers' scores on two separate, yet highly related, teacher efficacy instruments. Results suggest that the IC Teams intervention has an effect on

teachers' beliefs regarding how much they can do to increase the achievement of students who are considered disadvantaged for a variety of reasons and that following the intervention teachers in the treatment condition had increased efficacy for instructional strategies.

As stated above, the IC Teams intervention is a consultative approach in which every teacher in the school may not participate. Within the treatment schools in this study, an average of 65 percent of regular education teachers directly participated in the IC Teams intervention process. Therefore, the significant treatment effects found are encouraging. If more teachers participate in the intervention over time, larger effects may be found.

Relationship Between the Teacher Efficacy Scales

Correlations among unreliability-corrected school means were higher than the uncorrected school mean correlations. However, these correlations were not perfect. Therefore, this study provides minimal evidence that the TISA and TSES Efficacy for Instructional Strategies scales are measuring different aspects of teacher efficacy.

Relationships Between Efficacy Scales and Predictors

Secondary results indicate that male teachers scored higher than female teachers on the TISA (locus of control type teacher efficacy questions). Previous research suggests female teachers tend to have higher levels of teacher efficacy (Cheung, 2006). Therefore, this result is unanticipated and should be investigated further. The length of time a teacher has taught at a particular school was significantly positively related to scores on the TSES Efficacy for Instructional Strategies. This is an intuitively sensible finding as

teachers who have been teaching longer would presumably be more efficacious regarding their teaching strategies.

Limitations

Because of the design of the IC Teams program, all staff members in treatment schools may not be affected by the intervention in the first few years of its implementation. While there were significant effects for both outcomes, the effect sizes were relatively small. At this point in the evaluation process, all teachers have not interacted with IC Teams and therefore their belief systems may not yet have been influenced. For this reason, it will be important to continue the evaluation of the effect of IC Teams on teacher efficacy in the future years of the study. In addition, evaluation of the effects on those teachers who have had an IC case may be needed to determine if efficacy enhancement occurred specifically for those teachers.

Differential attrition could be a concern for this evaluation because of its possible impact on internal validity. While the overall attrition rate of 41% is considerable, it is almost all attributable to turnover, i.e., teachers leaving the school. This level of turnover for a three-year period is typical of schools (Marvel, Lyter, Pettola, Strizek, & Morton, 2007). Further, the results from attrition analyses suggest that the treatment conditions, teacher demographics, and pretest did not predict attrition. Therefore, the attrition threat to the internal validity of this evaluation seems low. Despite this, the supplementary complete case (listwise deletion) analysis which produces unbiased estimates only if data are missing completely at random (MCAR) produced similar, but not identical to results as the analysis that used the EM algorithm to impute values for cases with missing values. This method of imputation assumes that missing values are missing at random

(MAR). Importantly, the treatment effect was a significant predictor of both outcomes when cases with missing values were dropped or when values were imputed. The estimated size of the effect differs for the analyses using imputation and the complete case analysis for one of the outcome variables, however. Because of the desire to use an *intent to treat* approach, the moderately high level of missing values, and the non-significant relationship between treatment condition and missing outcome values, the results following imputation may be less biased if data are missing at random (Little & Rubin, 2002), but it is not possible to tell if this assumption is warranted.

Restriction of range of the outcome variables is also a concern as the mean scores on both teacher efficacy measures are relatively high. Whether this stems from weaknesses of the actual scales or that teachers in the selected schools really have high teacher efficacy scores, the restriction of range may attenuate the effect of IC Teams on teacher efficacy. Because significant effects were found, however, this concern is less problematic. In addition, both scales had relatively high reliabilities further suggesting that the score ranges were not too restricted.

Generalizability of results is also of concern as the results of this study may not apply to other school systems with different student and teacher demographics. The school system under investigation for this study volunteered to be part of a large research project, which may influence results in a positive way because the school administration is supportive of the intervention. However, it is important to note that the demographic make-up of the teachers within the school system is similar to the general population of elementary school teachers in the United States.

Finally, the TISA scale is newly developed and therefore construct validity is a concern. The relationship between the TISA scale other measures that fall in the spectrum of teacher locus of control/ general teacher efficacy should be investigated in future studies.

Future Research

Subsequent investigations of the effect of IC Teams on teacher efficacy should evaluate how the intervention influences efficacy beliefs through Bandura's four sources of efficacy. For example, investigating teachers' experiences, via interviews or other methods, regarding the specific components of IC Teams and how these experiences directly influence the sources of efficacy, could clarify the belief change process. Future investigations may evaluate the effect of IC Teams on collective teacher efficacy, which is defined as whether teachers believe the efforts of the school staff members as a whole will positively affect students (Goddard, Hoy, & Woolfolk Hoy, 2000). Altering the efficacy instruments to be more domain specific (i.e., targeted directly to efficacy for science instruction) as suggested by Bandura (1997) may also be explored to see if greater effects of the intervention can be detected.

Bandura (1997) suggested the most effect pathway for changes in efficacy are through mastery experiences. Therefore, teachers who have direct contact with IC Teams would be those who are most likely to undergo changes in efficacy. Future investigations may evaluate the efficacy outcomes for the particular teachers who directly participated in the IC Teams intervention. Future research should focus on how teacher efficacy is related to student achievement and how changes in teacher efficacy affect changes in student achievement. School achievement data are being collected as part of the larger

longitudinal evaluation and studies examining this relationship are in progress. It is important to determine the causal order of events regarding the implementation of IC Teams, changes in teacher efficacy, and changes in student achievement. For example, understanding whether changes in teacher efficacy lead to changes in student achievement or if the relationship is reversed should be considered as it is possible that a reciprocal relationship between the two variables exists. At this point, both the theory underlying the expectation that IC Teams will influence teacher efficacy and the present results are ambiguous with respect to the *process* through which such an effect might come about. For instance, the developers of IC Teams seem to expect that the consultation process and collegial assistance may directly lead to a greater sense of efficacy. On the other hand, substantial increases in teacher sense of efficacy may only follow successful experiences teachers have in serving difficult children. Further research that attempts to explore mediation and possible nonrecursive mechanisms in the causal process may be helpful. Evaluation of other student-level covariates and how they affect teacher efficacy at both the teacher and school levels is another area of potential investigation.

To gain a clearer understanding of the measurement issues regarding the TSES Efficacy for Instructional Strategies scales and the TISA, continued evaluation of these data is warranted. Utilizing the method described by Hox (2002), a multilevel factor model could be constructed by using the multivariate approach discussed in this paper to consider all 16 teacher efficacy questions as outcome variables. From this model, a covariance matrix of the 16 values could be created and input into a standard structural equation modeling program for further analysis.

Finally, although Likert scales are an easy and efficient way to measure many psychological variables and have a useful track record in attitude and personality research, these simple scales have many drawbacks. Exploration of alternative models that do not assume that the Likert-type items have interval properties and take their polytomous nature into account may provide fruitful in the future. One model that has been suggested to enhance the traditional Likert-response scale is to conduct a two-stage model that uses an unfolding technique as proposed by Albaum (1997). With this model, the respondent is first asked a traditional Likert-type question and then asked a *strength of feeling* question. Another option is to construct a continuum range of choices via the method described by Hodge and Gillespie (2003). With this continuum method, a phrase completion statement could be used at each anchor point. This method would have some advantages over the traditional Likert-type questions by being more analogous to a continuous response; thereby being more appropriate to use to meet the assumptions of parametric statistics. While future research might take a more analytical approach to the scaling of sense of efficacy, such an approach is beyond the scope of the present analyses

Implications

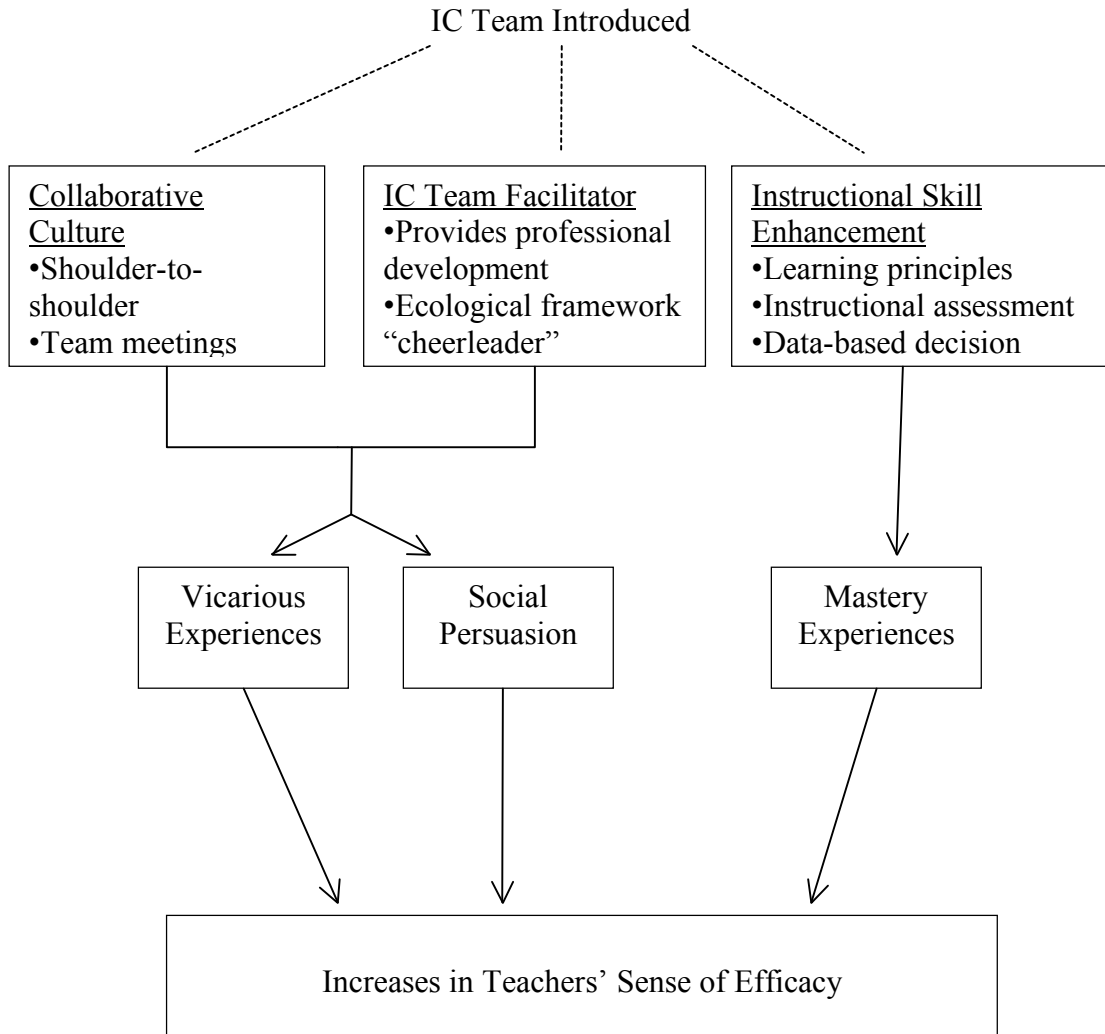
Although numerous studies have investigated teacher efficacy, efforts to fully understand this construct have proven elusive. Teacher efficacy should be studied using sound scales, methodologies, and statistics. This study attempted to address each of these three historical shortcomings in teacher efficacy research. Results from this study are useful for two reasons as this research has both a measurement and an intervention component.

For more than thirty years, teacher efficacy has been an important construct throughout educational research studies. Most of these studies have utilized an approach looking at two types of teacher efficacy. The first type, personal teacher efficacy, refers to a teacher's belief in his or her ability to have an effect on student learning, while general teaching efficacy is more closely aligned to the concept of locus of control, which assesses a teacher's belief regarding whether he or she can influence student learning despite any external constraints. From a psychometric perspective, this study provides limited support for the continued theoretical dichotomy of the teacher efficacy construct.

In terms of the influence of IC Teams on teacher efficacy, the significant results can be viewed as a meaningful advance regarding how teacher efficacy beliefs can be altered. IC Teams, as a comprehensive consultation package, may become a guide for intervention development targeted towards teacher belief change. The results of this study substantiate the results found by Vu et al. (2009) and provide a more detailed evaluation of the effect of IC Teams on teacher efficacy through two conceptual lenses.

Taken as a whole, the results of this study fill a gap in the teacher efficacy literature by providing evidence about the effect of a consultation-based intervention on teacher efficacy, while using a randomized design and a multivariate, multilevel analytic technique. This is the first known study of this kind to use this research design and methodology when evaluating the effects of an intervention on teacher efficacy.

Appendix A: Theoretical Model of Teacher Efficacy Change



Appendix B: Teacher Sense of Efficacy Scale

Efficacy for Instructional Strategies Scale

1. How well can you implement alternative teaching strategies in your classroom?
2. To what extent can you use a variety of assessment strategies?
3. To what extent can you provide an alternative explanation or example when students are confused?
4. To what extent can you craft good questions for your students?
5. How much can you do to adjust lessons to the proper level for individual students?
6. To what extent can you gauge student comprehension of what you have taught?

Note. Items 5 and 6 are not part of the original TSES - Short Form

Appendix C: Teacher Influence on Student Achievement Scale

1. How much can you do to increase the achievement of a student who has a specific learning disability?
2. How much can you do to increase the academic achievement of a student whose parents have a limited educational background?
3. If a student comes to your classroom lacking skills that should have been taught by previous teachers, how much can you do to help this child “catch up” in your classroom?
4. How much can you do to increase the achievement of a student from a disadvantaged family background?
5. To what extent can you help promote learning in students who receive special education services?
6. How much can you do to improve the academic performance of a student whose home environment lacks structure and discipline?
7. How much can you do in your classroom to improve the learning of a student with emotional and/or behavioral problems?
8. Within your classroom, how much can you help English Language Learners (ELL) improve their academic performance?
9. If a student in your class has parents who are not involved in the academic process, how much can you do to help this child learn?
10. How much can you do to increase the achievement of a student with attention problems?

Author Note

Jessica Koehler, Counseling and Personnel Services, University of Maryland.

This research is supported in part by grant number R305F050051 from the Institute of Education Sciences, U.S. Department of Education. Opinions expressed do not necessarily reflect the opinions or policies of the U.S. Department of Education. The development of the data used in this research would not have been possible without the members of the Experimental Evaluation of Instructional Consultation Teams research group at the University of Maryland (Sylvia Rosenfield, Gary D. Gottfredson, Amy Silverman, Todd Gravois, Deborah Nelson, Lauren Kaiser, Samantha Sedlik, Kate Bruckman, Phuong Vu, and Michael Rocque). A special thanks is also given to Jennifer Cassata and Kathy Aux (Prince William County Schools) for their help and support in collecting the data used for this study. Additionally, I would like to thank the technical support provided by Scientific Software International and Richard Tate (retired Florida State professor) for their help and guidance in creating the three-level, multivariate model in HLM6.

Correspondence concerning this manuscript should be addressed to Jessica Koehler, Counseling and Personnel Services, University of Maryland – College Park, College Park, MD 20742. Email: jes17@umd.edu.

Table 1

Factor Loadings for TISA^a and TSES^b Efficacy for Instructional Strategies Items

Item	Factor 1	Factor 2	<i>h</i> ²
How much can you do to improve the academic performance of a student whose home environment lacks structure and discipline?	.77	.19	.63
How much can you do to increase the achievement of a student from a disadvantaged family background?	.74	.30	.64
How much can you do to increase the academic achievement of a student whose parents have a limited educational background?	.71	.30	.60
How much can you do in your classroom to improve the learning of a student with emotional and/or behavioral problems?	.69	.22	.53
How much can you do to increase the achievement of a student with attention problems?	.69	.27	.55
If a student in your class has parents who are not involved in the academic process, how much can you do to help this child learn?	.69	.22	.52
How much can you do to “catch up” a student who come to our class reading two years below grade level?	.65	.29	.50
How much can you do in your classroom to increase the achievement of a student who has a specific learning disability?	.60	.42	.54
Within your classroom, to what extent can you help promote learning in students who receive special education services?	.60	.35	.49
Within your classroom, how much can you help English Language Learners improve their academic performance?	.52	.28	.34
To what extent can you use a variety of assessment strategies?	.22	.74	.58
How well can you implement alternative teaching strategies in your classroom?	.33	.70	.60
To what extent can you provide an alternative explanation or	.22	.67	.50

example when students are confused?			
To what extent can you craft good questions for your students?	.21	.64	.44
To what extent can you gauge student comprehension of what you have taught?	.28	.62	.45
How much can you do to adjust lessons to the proper level for individual students?	.37	.55	.43

Note: Principle factor analysis with varimax rotation. Loadings larger than .5 shown in boldface type. ^aTeacher Influence on Student Achievement. ^bTeacher Sense of Efficacy Scale.

Table 2

Level of Implementation and Utilization for Treatment Schools

School Code	Implementation quality (%)	Utilization (%)
58	<i>Missing Data</i>	<i>Missing Data</i>
60	95	50
79	88	34
66	<i>Missing Data</i>	<i>Missing Data</i>
67	100	78
2	100	50
97	98	56
86	98	60
20	90	61
82	99	65
4	100	67
15	95	67
43	99	70
33	92	74
13	99	74
74	96	80
23	99	93

Note. From Berger, Vaganek, Yiu, Hong, & Nelson (2009).

Table 3

Descriptive Statistics for the TISA^a and TSES^b Efficacy for Instructional Strategies

Outcome Variable	Treatment		Control	
	Mean	SD	Mean	SD
TISA scale	3.88	.49	3.78	.52
TSES EIS ¹	4.38	.42	4.31	.43

Note: n = 563 for treatment group, n = 640 for control group. ^aTeacher Influence on Student Achievement. ^bTeacher Sense of Efficacy Scale.

Table 4

Teacher-level Predictor Variables

Variable	Coding (if dummy variable)	Mean	SD
Years at school	0 < six years	.30	.46
	1 ≥ six years		
Job type ¹	0 regular education teacher	.38	.49
	1 non-regular education teacher ^a		
Educational level	0 < than a master's degree	.52	.50
	1 master's degree or higher		
Gender	0 female	.08	.27
	1 male		
Race/ethnicity	0 White, non-Hispanic	.13	.34
	1 Non-white		
Years of experience	0 < six years	.68	.47
	1 ≥ six years		
Age in years		41.13	11.90

Note: n = 1203 elementary school teachers. ^aTeachers were special education teachers or specialists.

Table 5

Correlations Among Teacher-level Variables

	TSES ^a <i>n</i> = 1203	TISA ^b <i>n</i> = 1203
TE Pretest	-.57**	.60**
Age in years	-.07*	-.03
Job type	-.04	.07*
Educational level	-.07*	.08**
Gender	-.03	.06*
Race/ethnicity	.00	.01
Years of experience	-.11**	.09
Years at school	-.07*	.00

Note: ^aRefers to the reflect and square root transformed TSES Efficacy for Instructional Strategies scale. ^bRefers to the Teacher Influence on Student Achievement scale.

* $p < .05$. ** $p < .01$.

Table 6

Correlations Between Scales at Each Level

<u>Teacher level</u>		<u>School level</u>	
SPSS 17 ^a	HLM 6	SPSS 17 ^b	HLM 6
0.66	0.66	0.67	0.81

Note: ^aBased on teacher-level observations, ignoring school membership. ^bBased on school means.

Table 7

Final Within School Model, Grand Mean Centered with Fixed Effects

<u>Fixed effects</u>			
Variable	Coefficient	Standard Error	<i>p</i>
TISA ^a Intercept, β_{10k}	.02	.03	.51
Pretest, β_{11k}	.59	.02	.00
Gender, β_{12k}	.31	.07	.00
TSES EIS ^b Intercept, β_{20k}	-.00	.03	.87
Pretest, β_{21k}	-.56	.02	.00
Years at school, β_{22k}	-.14	.04	.00
Random effects Level-2	Variance Component	<i>df</i>	<i>p</i>
TISA, r_{1jk}	.58	1167	.00
TSES EIS, r_{2jk}	.67	1167	.00
Random effects Level-2	Variance Component	<i>df</i>	<i>p</i>
TISA, u_{10k}	.02	33	.00
TSES EIS, u_{20k}	.01	33	.04

Note: Experimental treatment is coded as 1 = treatment, 0 = control. Pretest is the teacher efficacy measure for the baseline year. ^aTeacher Influence on Student Achievement. ^bTeacher Sense of Efficacy Scale, Efficacy for Instructional Strategies.

Table 8

Final Between School Model

<u>Fixed effects</u>			
<u>Variable</u>	<u>Coefficient</u>	<u>Standard Error</u>	<u>p</u>
TISA ^a Intercept, γ_{100}	-.08	.04	.04
Treatment, γ_{101}	.20	.06	.00
Pretest, β_{11k}	.59	.02	.00
Gender, β_{12k}	.31	.07	.00
TSES EIS ^b Intercept, γ_{200}	.07	.04	.05
Treatment, γ_{201}	-.16	.05	.00
Pretest, β_{21k}	-.56	.02	.00
Years at school, β_{22k}	-.14	.04	.00
Random effects Level-2	Variance Component	<i>Df</i>	<i>p</i>
TISA, r_{1jk}	.58	1167	.00
TSES EIS, r_{2jk}	.67	1167	.00
Random effects Level-3	Variance Component	<i>Df</i>	<i>p</i>
TISA, u_{10k}	.01	32	.01
TSES EIS, u_{20k}	.00	32	.22

Note: Experimental treatment is coded as 1 = treatment, 0 = control. Pretest is the teacher efficacy measure for the baseline year. ^aTeacher Influence on Student Achievement. ^bTeacher Sense of Efficacy Scale, Efficacy for Instructional Strategies.

Table 9

Variance Components and Proportion of Variance Explained

	Outcome Variable	
	TISA ^a	TSES EIS ^b
<i>Unconditional Model:</i>		
Level-2 ^c variance component	.93	.98
Level-3 ^d variance component	.03	.02
<i>Final Model</i>		
Level-2 variance component	.59	.67
Level-3 variance component	.01	.00
<i>Variance Explained (Percent)</i>		
Level-2 variance explained	37%	32%
Level-3 variance explained	68%	94%

Note: ^aTeacher Influence on Student Achievement Scale. ^bTSES Efficacy for Instructional Strategies scale.

^cWithin-school variance. ^dBetween-school variance.

Table 10

Final Between School Model Dropping Cases with Missing Data

<u>Fixed effects</u>			
Variable	Coefficient	Standard Error	<i>p</i>
TISA ^a Intercept, γ_{100}	-.08	.06	.19
Treatment, γ_{101}	.19	.09	.04
Pretest, β_{11k}	1.10	.03	.00
TSES EIS ^b Intercept, γ_{200}	2.24	.01	.00
Treatment, γ_{201}	-.03	.02	.06
Pretest, β_{21k}	-.24	.01	.00
Random effects Level-2	Variance Component	<i>df</i>	<i>p</i>
TISA, r_{1jk}	.72	660	.00
TSES EIS, r_{2jk}	.04	660	.00
Random effects Level-2	Variance Component	<i>df</i>	<i>p</i>
TISA, u_{10k}	.03	32	.00
TSES EIS, u_{20k}	.00	32	.10

Note: ^aTeacher Influence on Student Achievement. ^bTeacher Sense of Efficacy Scale, Efficacy for Instructional Strategies. Experimental treatment is coded as 1 = treatment, 0 = control. Pretest is the teacher efficacy measure for the baseline year.

Table 11

Descriptive Statistics for the TISA^a and TSES^b Efficacy for Instructional Strategies

Outcome Variable	Treatment		Control	
	Mean	SD	Mean	SD
TISA scale	3.87	.55	3.76	.60
TSES EIS ¹	4.37	.50	4.30	.51

Note: Descriptive statistics for data set when cases with missing values were deleted. $n = 338$ for treatment group, $n = 357$ for control group. ^aTeacher Influence on Student Achievement. ^bTeacher Sense of Efficacy Scale.

References

- Albaum, G. (1997). The Likert Scale revisited: An alternate version. *Journal of the Market Research Society, 39*, 331-349.
- Alkon, A., Ramler, M., & MacLennan, K. (2003). Evaluation of mental health consultation in child care centers. *Early Childhood Education Journal, 31*, 91-99.
- Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education, 17*, 86-95.
- Allinder, R. M. (1995). An examination of the relationship between teacher efficacy and curriculum-based measurement and student achievement. *Remedial and Special Education, 16*, 247-254.
- Alvidrez, J., & Weinstein, R. S. (1999). Early teacher perceptions and later student academic achievement. *Journal of Educational Psychology, 91*, 731-746.
- Anderson, R., Greene, M., & Loewen, P. (1988). Relationships among teachers' and students' thinking skills, sense of efficacy, and student achievement. *Alberta Journal of Educational Research, 34*, 148-165.
- Armor, D., Conroy-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A. (1976). *Analysis of the school preferred reading programs in selected Los Angeles minority schools* (No. R-2007-LAUSD). Santa Monica, CA: RAND. Retrieved January 8, 2008, from <http://www.rand.org/pubs/reports/R2007/>.
- Ashton, P. T. (1985). Motivation and teachers' sense of efficacy. In C. Ames & R. Ames (Eds.). *Research on motivation in education: Vol. 2. The classroom milieu* (pp. 141-174). Orlando, FL: Academic Press.

- Ashton, P. T. & Webb, R. B. (1982). *Teachers' sense of efficacy: Toward an ecological model*. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.
- Ashton, P. T., Buhr, D., & Crocker, L. (1984). Teacher's sense of efficacy: A self- or norm-referenced construct? *Florida Journal of Educational Research*, 26, 29-41.
- Ashton, P. T., & Webb, R. B. (1985). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.
- Bandura, A. (1977). Self-efficacy theory: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory* (Vol. Englewood-Cliffs, NJ): Prentice-Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Cambridge University Press.
- Berger, J., Vaganek, M., Yiu, E., Hong, V., & Nelson, D. (2009). *Treatment integrity and spread of use of Instructional Consultation Teams*. Manuscript in preparation.
- Berman, P., & McLaughlin, M.,W. (1977). *Federal programs supporting educational change; Vol. VII. Factors affecting implementation and continuation*, Santa Monica, CA: The RAND Corporation, Retrieved January 8, 2008, from <http://www.rand.org/pubs/reports/R1589.7/>.
- Berman, P., McLaughlin, M., Bass, G., Pauly, E., & Zellman, G. (1977). *Federal programs supporting educational change: Vol. VII. Factors affecting implementation and continuation* (Rep. No. R-1589/7-HEW). Santa Monica, CA: RAND. (ERIC Document Reproduction Service No. 140 432)

- Betz, N.E. & Schifano, R.S. (2000). Evaluation of an intervention to increase realistic self-efficacy and interests in college women. *Journal of Vocational Behavior*, 56, 35-52.
- Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, 75, 631-661.
- Brouwers, A., & Tomic, W. (2003). A test of the factorial validity of the Teacher Efficacy Scale. *Research in Education*, 69, 67-79.
- Cheung, Hoi-yan. (2006). The Measurement of Teacher Efficacy: Hong Kong Primary In-Service Teachers, *Journal of Education for Teaching: International Research and Pedagogy*, 32, 435-451.
- Coldarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education*, 60, 323-337.
- Cooper, H. M., Findley, M., & Good, T. (1982). Relations between student achievement and various indexes of teacher expectations. *Journal of Educational Psychology*, 74, 577-589.
- Cooper, H. M., & Tom, D. Y. (1984). Teacher expectation research: A review with implications for classroom instruction. *Elementary School Journal*, 85, 77-89.
- Costa, A. L. & Garmston, R. J. (1994). *Cognitive coaching: A foundation for renaissance schools*. Norwood, MA: Christopher-Cordon Publishers, Inc.
- Costas, L., Rosenfield, S., Gravois, T. (2003, August). *Impact of instructional consultation on teachers' satisfaction and skill development*. Poster session presented at the presented at the Annual American Psychological Convention, San Francisco, CA.

- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research, and evaluation, 10*(7), Available online: <http://pareonline.net/getvn.asp?v=10&n=17>.
- Cousins, J. B., & Walker, C. A. (2000). Predictors of Educators' Valuing of Systematic Inquiry in Schools. *Canadian Journal of Program Evaluation, Special Edition, 25-52*.
- Crano, W. D., & Mellon, P. M. (1978). Causal influence of teachers' expectations on children's academic performance: A cross-lagged panel analysis, *Journal of Educational Psychology, 70*, 39-49.
- Dembo, M. H., and Gibson, S. (1985). Teacher's sense of efficacy: An important factor in school improvement. *Elementary School Journal, 86*, 173-184.
- Deemer, S. A., & Minke, K. M. (1999). An investigation of the factor structure of the teacher efficacy scale. *Journal of Educational Research, 93*, 3-10.
- Deaver, J. (2005). A model of collaborative consultation to raise teacher self-efficacy. (Doctoral Dissertation, University of Colorado at Denver, 2005(Proquest No. AAT 3184197).
- DeForest, P. A., & Hughes, J. N. (1992). Effect of teacher involvement and teacher self-efficacy on ratings of consultant effectiveness and intervention acceptability. *Journal of Educational and Psychological Consultation, 3*, 301-316.
- Denzine, G. M., Cooney, J. B., & McKenzie, R. (2005). Confirmatory factor analysis of the Teacher Efficacy Scale for prospective teachers. *British Journal of Educational Psychology, 75*, 689-708.

- Edwards, J. L., Green, K. E., Lyons, C. A., Rogers, M. S., & Swords, M. E. (1998, April). The effects of cognitive coaching and nonverbal classroom management on teacher efficacy and perceptions of school culture. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling, 8*, 128-141.
- Emmer, E., & Hickman, J. (1991). Teacher efficacy in classroom management and discipline. *Educational and Psychological Measurement, 51*, 755-765.
- Evans, E. D., & Tribble, M. (1986). Perceived teaching problems, self-efficacy and commitment to teaching among preservice teachers. *Journal of Educational Research, 80*, 81-85.
- Firestone, W. A., & Rosenblum, S. (1988). Building commitment in urban schools. *Educational Evaluation and Policy Analysis, 10*, 285-300.
- Fives, H. & Alexander, P. A. (2004). Modeling teachers' efficacy, knowledge, and pedagogical beliefs. Paper presented at the Annual Meeting of the American Psychological Association Annual Meeting, Honolulu, HI.
- Fritz, Miller-Heyl, Kreutzer, and MacPhee (1995). Fostering personal teacher efficacy through staff development and classroom activities. *Journal of Educational Research, 88*, 200-208.
- Fox, M. E. (2004). School climate, teacher efficacy, and teachers' inclination to participate in consultation and referral (Doctoral dissertation, Fordham University, 2004). (Proquest No. AAT 3134437).
- Fritz, J. J., Miller-Heyl, J., Kreutzer, J. C., & MacPhee, D. (1995). Fostering personal

- teaching efficacy through staff development and classroom activities. *Journal of Educational Research*, 88, 200-208.
- Fudell, R. (1992). Level of implementation of teacher support teams and teachers' attitudes toward special needs students. *Dissertation Abstracts International*, 53(05), 1399A (University of Microfilms No. AAC-922763).
- Fullan, M. (1991). *The new meaning of educational change*. New York, NY: Teachers College Press.
- Fuller, F. (1969). Concerns of teachers. *American Educational Research Journal*, 6, 207-227.
- Geller, S., & Lynch, K. (1999). *Teacher Opinion Survey*. Richmond: Virginia Commonwealth University Intellectual Property Foundation and Wingspan, LLC.
- Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: a construct validation. *Journal of Educational Psychology*, 76, 569-582.
- Goddard, R. D., Hoy, W. K., & Woolfolk Hoy, A. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal*, 37, 479-508.
- Gravois, T., Fudell, R., & Rosenfield, S. (2005). *The Level of Implementation Scale for Instructional Consultation Teams: Administration & Scoring Guide*. [Unpublished manual], College Park, MD.
- Gravois, T., Rosenfield, S., & Gickling, E. (2002). *Instructional consultation teams: Training manual*. University of Maryland, Instructional Consultation Lab, College Park, MD.
- Gravois, T. A., & Rosenfield, S. (2002). A multidimensional framework for evaluation of

- instructional consultation teams. *Journal of Applied School Psychology, 19*, 5-29.
- Gravois, T. A., & Rosenfield, S. (2006). Impact of Instructional Consultation Teams on the disproportionate referral and placement of minority students in special education. *Remedial and Special Education, 27*, 42-52.
- Grinder, M. (1996). *Envoy: Your personal guide to classroom management*. Battle Ground, WA: Michael Grinder Associates.
- Greenwood, G. E., Oljenik, S. F., & Parkay, F. W. (1990). Relationship between four teacher efficacy belief patterns and selected teacher characteristics. *Journal of Research and Development in Education, 23*, 102-106.
- Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review, 15*, 147-179.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley & Sons.
- Guskey, T. R. (1981). Measurement of responsibility teachers assume for academic successes and failures in the classroom. *Journal of Teacher Education, 32*, 44-51.
- Guskey, T. R. (1984). The influence of change in instructional effectiveness upon affective characteristics of teachers. *Educational Researcher, 81*, 41-47.
- Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal, 13*(3), 627-643.
- Hans, T. (2000). A meta-analysis of the effects of adventure programming on locus of control. *Journal of Contemporary Psychotherapy, 30*, 33-60.
- Hattie, J. A., Marsh, H. W., Neill, J. T. & Richards, G. E. (1997). Adventure Education and Outward Bound: Out-of-class experiences that have a lasting effect. *Review of*

- Educational Research*, 67, 43-87.
- Henson, R. K. (2002). From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist*, 37, 137-150.
- Henson, R. K., Kogan, L. R., & Vacha-Haase, T. (2001). A reliability generalization study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement*, 61, 404-420.
- Hodge, D. R. and Gillespie, D. (2003). Phrase Completions: An Alternative to Likert Scales. *Social Work Research*, 27, 45-55.
- Hoover-Dempsey, K. V., Bassler, O. C., & Brissie, J. S. (1987). Parent involvement: Contributions of teacher efficacy, school socioeconomic status, and other school characteristics. *American Educational Research Journal*, 85, 287-294.
- Hox, J. (2002). *Multilevel Analysis*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Hoy, W. K., & Woolfolk, A. E. (1993). Teachers' sense of efficacy and the organizational health of schools. *The Elementary School Journal*, 93, 355-372.
- Huang, X., Liu, M., & Shiomi, K. (2007). An analysis of the relationships between teacher efficacy, teacher self-esteem and orientations to seeking help. *Social Behavior and Personality*, 35, 707-716.
- Knotek, S. E., Rosenfield, S. A., Gravois, T. A., & Babinski, L. M. (2003). The process of fostering consultee development during instructional consultation. *Journal of Educational and Psychological Consultation*, 14, 303-328.
- Krpn, M. M. (1997). Cognitive coaching and efficacy, growth, and change for second-, third-, and fourth-year elementary school educators. (Doctoral Dissertation,

- California State University, Fullerton, 1997). (ProQuest No. AAT 1384152).
- Kushner, S. (February, 1993). *Teacher efficacy and preservice teachers: A construct validation*. Paper presented at the Paper presented at the Annual Meeting of the Eastern Educational Research Association., Clearwater Beach, FL.
- Levinsohn, M. R. (2000). *Evaluating Instructional Consultation Teams for student reading achievement and special education outcomes*. Unpublished doctoral dissertation, University of Maryland, College Park.
- Little, R. & Rubin, D. B. (2002). *Statistical Analysis with Missing Data* (2nd ed.). New York: John Wiley.
- Marvel, J., Lyter, D. M., Pettola, P., Strizek, G. A., & Morton, B. A. (2007). *Schools and Staffing Survey, Teacher Attrition and Mobility: Results from the 2004-05 Teacher Follow-Up Survey*. (National Center for Education Statistics 2007-307). US Department of Education, Washington, DC. Retrieved May 4, 2009 from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007307>
- McKenna, S. A. (2005). Validation of the Instructional Consultation Teams Level of Implementation Scale-Revised, (Doctoral Dissertation, University of Maryland, College Park, 2005). (Retrieved from <http://hdl.handle.net/1903/2421>).
- Middeton, V. (2002). Increasing pre-service teachers' diversity beliefs and commitment. *The Urban Review*, 34, 343-361.
- Miller, A. D. & Murdock, T.B. (2007). Modeling latent true scores to determine the utility of aggregate student perceptions as classroom indicators in HLM: The case of classroom goal structures. *Contemporary Educational Psychology*, 32, 83-104.
- Midgley, C., Feldlaufer, H., & Eccles, J. (1989). Change in teacher efficacy and student

- self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, 81, 247-258.
- Parkay, F. W., Greenwood, G. E., Oljenik, S. F., & Proller, N. (1988). A study of the relationship among teacher efficacy, locus of control, and stress. *Journal of Research and Development in Education*, 21, 13-22.
- Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology*, 76(1), 85-97.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods, 2nd Edition*. Newbury Park, CA: Sage Publications.
- Reschly, D. J. (2008). School psychology, paradigm shift and beyond, In A. Thomas & J. Grimes (Eds.). *Best practices in school psychology V* (pp. 3-16). Washington, DC: National Association of School Psychologists.
- Rhem, J. (1999). Pygmalion in the classroom. *The National Teaching and Learning Forum*, 8, 1-4.
- Rimm-Kaufman, Sawyer, B. E. (2004). Primary-grade teachers' self efficacy beliefs, attitudes toward teaching, and discipline and teaching practice priorities in relation to the "Responsive Classroom" approach. *The Elementary School Journal*, 104, 321-341.
- Rose, J. S., & Medway, F. J. (1981). Measurement of teachers' beliefs in their control over student outcomes. *Journal of Educational Research*, 74, 185-191.
- Rosenfield, S. A. (1987). *Instructional Consultation*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

- Rosenfield, S. (2004). Consultation as dialogue: The right words at the right time. In N. Lambert, I. Hylander, & J. Sandoval. *Consultee-Centered Consultation: Improving the quality of services in schools and community organizations* (pp. 337-347). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rosenfield, S. (2008). Best practice in instructional consultation and instructional consultation teams, In A. Thomas & J. Grimes (Eds.). *Best practices in school psychology V* (pp. 1645-1660). Washington, DC: National Association of School Psychologists.
- Rosenfield, S. & Gottfredson, G. (2004). *Evaluating the Efficacy of Instructional Consultation Teams*. Unpublished grant proposal. University of Maryland, Department of Counseling and Personnel Services. Retrieved November 20, 2008 from <http://www.icteams.umd.edu/Proposal%20for%20Project.pdf>
- Rosenfield, S., Silva, A., & Gravois, T. (2008). Bringing instructional consultation to scale: Research and development of IC and IC Teams. In W. P. Erchul & S. M. Sheridan (Eds.). *Handbook of research in school consultation: Empirical foundations for the field* (203-223). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rosenfield, S., & Gravois, T. A. (1996). *Instructional consultation teams: Collaborating for change*. New York: Guilford Press.
- Ross, J. A. (1992). Teacher efficacy and the effects of coaching on student achievement. *Canadian Journal of Education*, 17, 51-65.
- Ross, J. A. (1994). Beliefs that make a difference: The origins and impacts of teacher efficacy. (Report No. ED 379 216). Calgary, Alberta, Canada.

- Ross, J. A. (1998). The antecedents and consequences of teacher efficacy. In J. Brophy (Ed.), *Research on teaching* (Vol. 7, pp. 49-74). Greenwich, CT: JAI Press.
- Ross, J. A. & Bruce, C. (2007). Professional development effects on teacher efficacy: Results of a randomized field trial. *Journal of Educational Research, 101*, 50-60.
- Rotter, J. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*(Whole No. 609).
- Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in personality and social psychology bulletin. *Personality and social psychology bulletin, 28*(12), 1629-1646.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.
- Schafer, J. L., & Graham, J. W. (2002). Missing Data: Our view of the state of the art. *Psychological Methods, 7*, 147-177.
- Smylie, M. A. (1988). The enhancement function of staff development: Organizational and psychological antecedents to individual change. *American Educational Research Journal, 25*, 1-30.
- Sniders, T. & Bosker, R. (1999) *Multilevel Analysis*. Sage Publications, London.
- Sofronoff, K. & Farbotko, M. (2002). The effectiveness of parent management training to increase self-efficacy in parents and children with Asperger syndrome. *Autism, 6*, 271-286.
- Soodak, L. C., & Podell, D. M. (1993). Teacher efficacy and student problem as factors in special education referral. *The Journal of Special Education, 27*, 66-81.

- Soodak, L. C., & Podell, D. M. (1994). Teachers' thinking about difficult-to-teach students. *Journal of Educational Research, 88*, 44-51.
- Soodak, L. C., & Podell, D. M. (1997). Teacher efficacy: Toward the understanding of a multi-faceted construct. *Teaching and Teacher Education, 12*, 401-412.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.
- Tate, R. L., & Pituch, K. A. (2007). Multivariate hierarchical linear modeling in randomized field experiments. *The Journal of Experimental Education, 75*, 317-337.
- Tilly, W. D. (2008). The evolution of school psychology to science-based practice: Problem-solving and the three-tiered model, In A. Thomas & J. Grimes (Eds.). *Best practices in school psychology V* (pp. 17-36). Washington, DC: National Association of School Psychologists.
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.
- Tschannen-Moran, M., Woolfol Hoy, A., & Hoy, W. K. (1998). Teacher Efficacy: Its Meaning and Measure. *Review of Educational Research, 68*, 202-248.
- Tournaki, N. & Podell, D. M. (2005). The impact of student characteristics on teachers' predictions of student success. *Teaching and Teacher Education, 21*, 299-314.
- Vu, P., Bruckman, K., Koehler, J., Kaiser, L., Rosenfield, S., Nelson, D., et al. (2009). *The effects of Instructional Consultation Teams on teacher beliefs and instructional practices after two years of implementation*. Manuscript submitted for publication.

Wigfield, A., Galper, A., Denton, K., & Seefeldt, C. (1999). Teachers' beliefs about former Head Start and non-Head Start first-grade children's motivation, performance, and future educational prospects. *Journal of Educational Psychology, 91*, 98-104.

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

Woolfolk, A. E., Rosoff, B., & Hoy, W. K. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and Teacher Education, 6*, 137-148.