ABSTRACT

Title of Document:	THE ORGANIZATIONAL EFFECT OF COLLECTIVE TEACHER EFFICACY: A STUDY OF STUDENT TEST SCORES AND HIGH SCHOOLS WITH LARGE CONCENTRATIONS OF MINORITY STUDENTS
	Jeehye Shim Deogracias, Doctor of Philosophy, 2012
Directed By:	Professor Robert G. Croninger Department of Teaching, Learning, Policy and Leadership

This study examined collective teacher efficacy along with student mathematics and reading test scores on a nationally representative sample of high school students and their schools. Collective teacher efficacy is defined as a group of teachers' shared belief in their ability to promote learning and positive student outcomes. For this study, this construct is conceptualized as an aspect of the informal organization of schools, whereas other factors, such as curriculum tracking, are considered to be an aspect of the formal organization of schools. Prior research into collective teacher efficacy as an organizational construct found evidence of a positive relationship with student achievement scores, though peer-reviewed studies have not been done on a national sample of students. In addition, there has been no research on the possible moderating effects of collective teacher efficacy.

I used a national dataset, the National Longitudinal Survey of 1988, and hierarchical linear modeling as the quantitative method. Contrary to prior research, I found no evidence that collective teacher efficacy had any effect on high school mathematics or reading test scores. It was not associated with either outcome, nor did it moderate the effect of the school's minority enrollment. Moreover, the largest predictor of high school test scores was prior achievement, which suggests that future research should examine school effects for young children. While this study confirmed the existence of an achievement gap between minority and majority students within schools, this gap did not vary between schools and thus, could not be modeled as a function of school characteristics. One school measure, academic press, had an impact only after controlling for average prior achievement. Additional efforts should be made to develop better measures of school organization, particularly the informal aspects of schooling, such as a school's academic press.

THE ORGANIZATIONAL EFFECT OF COLLECTIVE
TEACHER EFFICACY: A STUDY OF STUDENT TEST SCORES AND
HIGH SCHOOLS WITH LARGE CONCENTRATIONS OF MINORITY STUDENTS

by

Jeehye Shim Deogracias

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
2012

Advisory Committee:

 Dr. Robert G. Croninger, Chair, Associate Professor
 Dr. Thomas Davis, Assistant Professor
 Dr. Susan De La Paz, Dean's Representative, Associate Professor
 Dr. Sherick Hughes, Assistant Professor
 Dr. Jennifer King Rice, Professor

DEDICATION

To my family: past, present, and future

ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation to a number of individuals who were instrumental to the completion of this dissertation. First, I would like to thank my dissertation chair, Bob Croninger, for many years of advice, patience, and endless knowledge on all things education, analysis, and policy. I would also like to thank my committee members, Drs. Tom Davis, Susan De La Paz, Sherick Hughes, and Jennifer Rice for their time and wisdom on this project and on others; on educational subjects and beyond.

Second, countless friends and classmates have supported me through this journey. To Erin Ward Bibo and Laura Yee: Thank you for writing with me, supporting me at all hours, and laughing with me when it was much needed. I look forward to witnessing the amazing things that lie ahead for both of you.

Lastly, I would not be who I am and would not be passionate about what I do without my family. Thank you to my parents, Jay and Hyekyung Shim, for the many sacrifices that were made that allowed me to pursue my dreams and aspirations. This degree is just as much yours as it is mine. Many thanks to Michael Shim for studying with me this past year, renewing those factor analysis books 3,000 times at your library, and for making me a very lucky sister 28 years ago. To Karen and Frank Deogracias: Thank you for the multitude of ways you both have supported me through two graduate degrees and the many years of being included in your family. And finally to my husband, Andrew: Your unending love, support, and laughs got me through the five chapters of this manuscript and the many chapters in our lives thus far. Thank you for being my partner in all things; let the next chapter of our lives begin!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

We report to the American people that while we can take justifiable pride in what our schools and colleges have historically accomplished and contributed to the United States and the well-being of its people, the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people.
(A Nation at Risk, 1983)

Policymakers have long recognized that schools are important. The opening quote, cited from the 1983 government-sponsored report, *A Nation at Risk*, described the state of the United States education system and its impact on the future of the country and its economy. This report on education highlighted what many people believed to be the fundamental role of education in fostering the nation's economic prosperity and health. In addition to the economy, policymakers and researchers frequently cite schools as solutions to America's crime rates (Heckman & Krueger, 2005; Moretti, 2007), improvements in health (Muenning, 2007), and increased social mobility (Entwisle, Alexander, & Olson, 2005; McMurrer & Sawhill, 1998; Rouse & Barrow, 2006).

With increased focus, however, brings increased criticism. The opening quote, while describing the past accomplishments of our education system positively, delivers a scathing critique on the current state of education. The authors of *A Nation at Risk* (1983) cited data that found about 13 percent of 17-year old students in the 1980s were functionally illiterate and that the rate amongst minority students may have been as high as 40 percent. While students in 2009 scored higher than students in 1990 in

mathematics and higher than students in 1992 in reading, Black, Latino/a, and Native American/Alaska Native students consistently scored below White and Asian/Pacific Islander students (U.S. Department of Education, 2011). These results were consistent across multiple grade levels and throughout the respective time spans.

Researchers have investigated multiple ways to improve student learning and achievement through increasing teacher quality (e.g., Rice, 2003), school resources (e.g., Hanushek, 1997), parental involvement (e.g., Dearing, Kreider, Simpkins, & Weiss, 2006), and even through providing housing vouchers (e.g., Ladd & Ludwig, 2003). This study, however, investigates the ways in which certain school characteristics and organizational features impact student achievement – most notably, whether a collective sense of teacher efficacy promotes higher levels of student achievement.

Collective Teacher Efficacy

Albert Bandura is credited with developing the theory and conceptualization of both self-efficacy and collective efficacy. While both of these areas are theoretically similar, they are nonetheless conceptually different (Goddard, 2001; Goddard & Goddard, 2001; Tschannen-Moran & Barr, 2004). A teacher's sense of efficacy, or a teacher's belief in his or her ability to promote learning (Bandura, 1993), originated within social cognitive theory, which views human behavior as the "product of a dynamic interplay of personal, behavioral, and environmental influences" (Pajares, 2002, p. 1). However, collective teacher efficacy can also be seen as an organizational characteristic (Goddard & Goddard, 2001) and a property of schools (Goddard, 2001; Lee, Dedrick, & Smith, 1991; Tschannen-Moran & Barr, 2004).

Collective teacher efficacy is a relatively new area of research (Evans, 2009; Goddard & Goddard, 2001) and is increasingly becoming recognized as an important aspect of schooling. Collective teacher efficacy is the group's shared belief in its capability to organize, execute, and produce positive student outcomes (Goddard, 2001; Goddard & Goddard, 2001; Goddard, Hoy, & Woolfolk Hoy, 2000). Schools are an appropriate environment for research on collective efficacy because teachers operate collectively within this "interactive social system" (Bandura, 1993, p. 141) to meet organizational goals (Evans, 2009). In other words, the mutual dependencies associated with schools as an organization highlight the importance of understanding how a collective sense of efficacy influences student outcomes.

Collective teacher efficacy may influence a group's performance by shaping the behavioral and normative environment of the organization (Goddard, 2001; Goddard et al., 2000). Collective teacher efficacy is developed in a number of ways: through everyday interactions with students, fellow teachers, and administrators; through perceptions about their colleagues' competence in teaching; through cumulative teaching experiences; the availability of school resources; from feelings of control within the classroom and the school; and the demands associated with student characteristics (Bandura, 1993, 2000; Evans, 2009; Goddard, 2001; Lee et al., 1991). Once it is established, collective teacher efficacy is a relatively stable property that requires substantial effort to change (Goddard et al., 2000; Moore & Esselman, 1992).

Beliefs in a faculty's ability to teach their students can have both positive and negative effects on how well schools function as a social system (Bandura, 1993; Brookover et al., 1978). Researchers have found that schools with high collective teacher

efficacy promote productive teacher behaviors such as the acceptance of challenging goals, production of strong organizational effort, and a strong sense of persistence that lead to better student outcomes (Goddard, 2001; Goddard et al., 2000; Sweetland & Hoy, 2000). In some studies, the impact of collective teacher efficacy on student achievement is so large that the effect of collective teacher efficacy is greater in magnitude than the impact of any student demographic controls including student socioeconomic status (SES) (Bandura, 1993; Goddard et al., 2000). This result suggests that, on average, any negative impact that a student's SES may have on his/her achievement may be made up with high levels of collective teacher efficacy in their school. Furthermore, Tucker and colleagues (2005) found that efforts to increase teacher efficacy are vital in "increasing the low academic achievement and decreasing the disproportionate high school dropout rates among culturally diverse students" (p. 31).

This study builds upon prior research on collective teacher efficacy and its influence on student achievement. In addition, this study extends beyond examinations of schools in general by investigating the impact that collective teacher efficacy may have on minority students and schools.

The Effect of Minority Students and Schools

Examinations of students of color are an important area of education research. Black and Latino/a students are more likely to attend racially segregated and high poverty schools (Rumberger & Palardy, 2005). Not only do studies confirm the existence of an achievement gap between Black and White students, but this gap appears before children enter kindergarten and persists into adulthood (Jencks & Phillips, 1998).

One of the most influential education studies, the *Equality of Educational Opportunity* study by Coleman et al. (1966), found that student achievement was more related to student characteristics such as race and SES than school characteristics and resources. This study has been widely cited as proof that schools cannot counteract the effect of a student's background. However, Borman and Dowling (2010) recently published a reanalysis of the Coleman data using more sophisticated methodologies such as hierarchical linear modeling and regression analysis. These authors found that 40 percent of the variation in student achievement was between schools, even after controlling for student characteristics. In other words, differences in school resources and their social organization significantly impact individual student's learning. In addition, Borman and Dowling found that certain aspects of a school's social organization offered a more equitable distribution of learning within the school. They found that the achievement gap between Black and White students and the gap between students of high and low socioeconomic status was in part due to teachers' preferential biases towards middle class students and the consequences of curriculum tracking. These authors negate the original findings of the Coleman Report, and found that school characteristics were more influential than individual student's race/ethnicity and SES. To put it differently, schools have the power to improve or exacerbate the gap in learning among different types of students.[1]

---

[1] Borman and Dowling (2010) state, however, that the Coleman report did not include much data on veiled inequalities that may exacerbate inequalities within the schools. Research into the relationships between teachers and their students, the deployment of resources, and the quality of education may help explain the differences within and between segregated schools and more integrated schools.

Some suspect that racial attitudes, stereotypes, and perceived ability to work with minority students play a role in teachers' expectations and beliefs of teaching certain racial and ethnic groups (Diamond, Randolph, & Spillane, 2004; Jussim, Eccles, & Madon, 1996; Tucker et al., 2005). For example, using qualitative methods, Diamond, Randolph, and Spillane (2004) found that teachers in predominantly Black schools emphasized students' deficits and had a reduced sense of responsibility for their learning. In predominantly White or Asian schools, however, students' intellectual assets were emphasized, and teachers felt more accountable for what their students learned. These feelings remained, even in predominantly low-income Asian schools. Diamond and his colleagues found that teachers felt more positive about teaching Asian students because of positive stereotypes that they held about Asian students.

Other researchers suggest that students of different racial/ethnic backgrounds may react differently toward their teachers' beliefs and attitudes. Jussim, Eccles, and Madon's (1996) study of middle school students in Michigan found teacher expectations (as measured by teachers' perceptions of performance) influenced African American students more negatively and more strongly than White students. Similarly, Brookover and his colleagues (1978) found that collective feelings of teacher's commitment to doing a good job were impactful only in majority Black schools; these collective feelings had no impact in majority White schools.

Specific to collective teacher efficacy, current research in this area and its relationship to high minority population schools are mixed. Some researchers found no relationship (e.g., Goddard, 2001; Goddard, LoGerfo & Hoy, 2004; Goddard & Skrla, 2006) while other researchers found a statistically significant one (e.g., Newman et al.,

1989). The literature demonstrates the need for additional research into how schooling in high minority enrollment schools improves and/or restricts learning, especially, for the purpose of this study, in regards to collective teacher efficacy.

Purpose of the Study

The purpose of this study is to examine the relationship between collective teacher efficacy and student achievement for early high school students. In addition, this study investigates the possible moderating role collective teacher efficacy may have on the achievement of individual minority students within schools and on the average achievement of students in high schools with large concentrations of minority students. This study focuses on three research questions:

1. Does collective teacher efficacy have a positive impact on student achievement?

2. Does a gap exist between minority students and majority students in regards to achievement within schools? If so, do high schools with high collective teacher efficacy have an equalizing effect on this achievement gap within schools?

3. Do high schools with large minority concentrations have different achievement levels compared to schools with low minority concentrations? If so, do high schools with high collective teacher efficacy have an interacting or moderating effect on the relationship between schools with high minority populations and student achievement?

This study contributes to the literature on collective teacher efficacy in a number of ways. First, research on teacher efficacy typically deals with individual teachers as the unit of analysis. This study, however, views teacher efficacy as a collective school

property, thus bringing the unit of analysis to the school level. The relationship between collective teacher efficacy and student achievement is not well studied (Goddard & Goddard, 2001; Pajares, 1997). To date, only 12 peer-reviewed quantitative studies conceptualize and test for collective teacher efficacy in this way, though examinations into this construct has been increasingly popular in the last decade.

Second, much of the research on collective teacher efficacy uses small scale, local studies to examine this construct. This study uses a national dataset, the National Education Longitudinal Study of 1988 (NELS:88), which provides a number of distinct advantages. First, it allows for the generalization of results to a broader population of students in the United States. Although the data are roughly twenty years old, policy makers dealing with nationwide education policies may find it useful in determining the generic relationship between teacher efficacy and student outcomes. Second, the NELS:88 dataset has a very large sample size that facilitates statistical power and accuracy (Ware & Kitsantas, 2007). Additionally, NELS:88 not only surveyed students from across the country but also their parents, teachers, and school administrators. These multiple sources of data allow for a robust picture into the relationships and phenomenon that occur within schools. Lastly, NELS:88 followed up with the same students at multiple time points. The longitudinal nature of this data source allows a researcher to control for student knowledge and experiences at an early time point in order to examine the influence of collective teacher efficacy on student knowledge at a later time point.

A third contribution of this study to the literature involves its focus on important equity issues. While researchers have examined relationships between school minority composition, teacher efficacy, and student achievement, no studies have examined the

8

possibility of collective teacher efficacy moderating the relationship between student achievement and high minority population schools. This study aims to fill this gap in the literature on collective teacher efficacy. Lastly, this study also examines whether collective teacher efficacy plays an important role in improving the equitable distribution of outcomes in schools by exploring how this construct affects individual students based on their race and ethnicity.

Definitions of Terms

Teacher efficacy: A teacher's belief in his or her ability to promote learning and produce positive student outcomes

Collective teacher efficacy: As a group, teachers' shared belief in their ability to produce positive student outcomes

Minority: For the purposes of this study, "minority" refers to African Americans, Latino/as, American Indians, and Alaska Natives. All of these groups have their own unique identities, histories, and experiences and designating any group as "minority" blurs these distinctions and threatens to become a stereotype. However, I chose to combine these groups together because these groups have a history of marginalization and often have similar educational experiences.[2] Moreover, although NELS:88 is a national sample, there are insufficient sample sizes within schools to examine the distinct relationships for each racial/ethnic group.

---

[2] Data based on the National Assessment on Educational Progress from 1990 to 2009 show Black and Hispanic students do less well on mathematics and reading than White and Asian students (Department of Education, 2011).

9

Majority: For the purposes of this study, "majority" refers to non-Hispanic Whites, Asians, and Pacific Islanders. These groups are also unique in identity, history, and experience (see S. Lee, 1996 and Lew, 2006), and designating any group as "majority" blurs these distinctions and threatens to become a stereotype. However, I chose to combine these groups together because these two groups are similar in their educational experiences and outcomes, and the within school sample size is too small to model the distinct relationships for each racial/ethnic group.[3] Designating this group as "majority" does not imply a numerical majority. In some parts of the country, ethnic minority groups are becoming the numerical majority (Mellnik, 2012).

Achievement: For the purposes of this study, achievement is represented as test scores. I recognize that most test scores are not representative of true achievement and learning, in addition to the fact that they include measurement error and are vulnerable to corruption and inflation (Koretz, 2002). Test scores, whether they represent true achievement or not, are important in that so much of our education system relies on such numbers. Much of the quantitative literature on achievement relies solely on test scores, largely because no alternative indicators of achievement are included in national surveys.

Moderator: A variable that affects the strength or direction of a particular relationship (Baron & Kenny, 1986). A statistically significant interaction between a predictor and the moderator variable supports the existence of a moderating variable. See Figure 1.1.

---

[3] For example, Diamond, Randolph, and Spillane (2004) found differences in teacher's beliefs in their students' academic abilities depending on the racial makeup of the school: Teachers of schools with low income and African American students had a reduced sense of responsibility for their learning, whereas teachers of middle-income, predominantly White or Asian schools felt more accountable for what their students learned.

Figure 1.1: Moderator variable

Independent Variable → Dependent Variable

Moderator →

Independent Variable x Moderator →
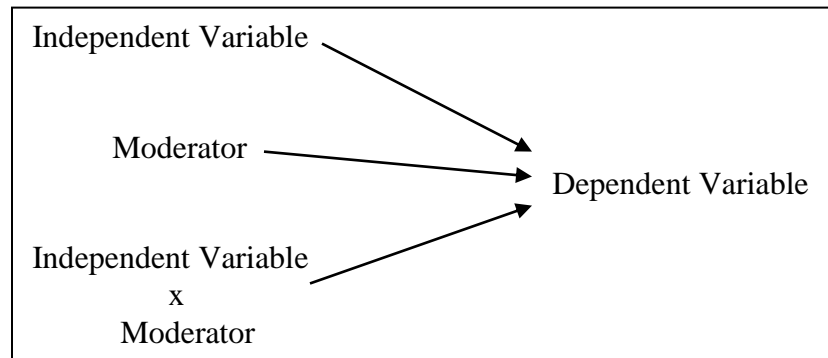
Conceptual Model

This study draws upon multiple theories and lenses. First, this study draws from

the effectiveness of schools literature or school-effects literature. School-effects

literature views school characteristics, such as structure and organizational properties, as

important factors that influence student outcomes (Lee, 2000; Lee & Bryk, 1989). A

sociological approach to school-effects literature is appropriate because schools are mini-

societies with unique environments, norms, practices, interactions, and individuals. As

Bidwell and Kasarda (1980) suggest, I conceptualize the school-effects literature as

having two distinctions, schools and schooling. Schools are organizations in which

instruction occurs, whereas schooling is the process by which instruction occurs.

Schools, adhering to this distinction, are typically thought of as resources, structural

characteristics and demographics. Some examples of schools include size, sector (e.g.,

public, Catholic), location, minority student concentration, average student SES, and

average teacher experience. Schools influence schooling in that schools provide the

11

setting for individuals, but the interactions and relationships within them are where schooling occurs.

Within the umbrella of schooling, I conceptualize schooling within organization theory, and examine it in two ways. The first is through a formal organizational lens. A formal organization lens of schooling aims to enhance the efficiency and effectiveness of schools through the formalization and standardization of tasks, specialization, authority structure, uniformity, and rationality in actions and behaviors (Hanson, 2003; Hoy & Miskel, 2008; Weber, 1909-1920). Examples of formal organizations of schooling include curricular tracking and subject area departments.

The process of schooling not only involves formal organizations but informal organizations as well. The informal organization of schooling is characterized by processes and behaviors that are not formally planned but emerge to fulfill a need (Hanson, 2003). Some examples of informal organizations of schooling include levels of collegiality among teachers, teachers' collective feelings about their principal's leadership, sense of control over their classroom, and the main focus of this study, their collective sense of efficacy.[4]

Figure 1.2 graphically illustrates the conceptual model. Consistent with school-effects literature, schools and schooling are distinct aspects. The process of schooling is divided into two components, formal and informal organization. The primary relationship of interest, which is addressed by research question 1, is between collective

---

[4] Current reforms, such as professional learning communities, are blurring the lines of formal and informal aspects of schooling. For example, some interventions launch programs to develop collegial relationships and increase decision making within schools. Since the survey I used in this study was initiated in 1988, the blurring of these distinctions was not much of a concern for schools in the late '80s/early '90s.

teacher efficacy and student achievement (A). If high minority enrollment schools have

differences in test scores (D), then research question 3 examines the way collective

teacher efficacy influences the relationship between minority school composition and

student outcomes (B). If minority and majority student have differences in test scores

(F), the ways in which collective teacher efficacy influences this gap in achievement will

be addressed in research question 2 (C). In order to account for alternative influences, I

will control for school characteristics (D), formal organization of schooling (E), and

informal organizational aspects of schooling (A). Lastly, to isolate the effect of schools

on students, I will control for students' demographics and prior achievement levels as

well (F).

Figure 1.2: Conceptual Model
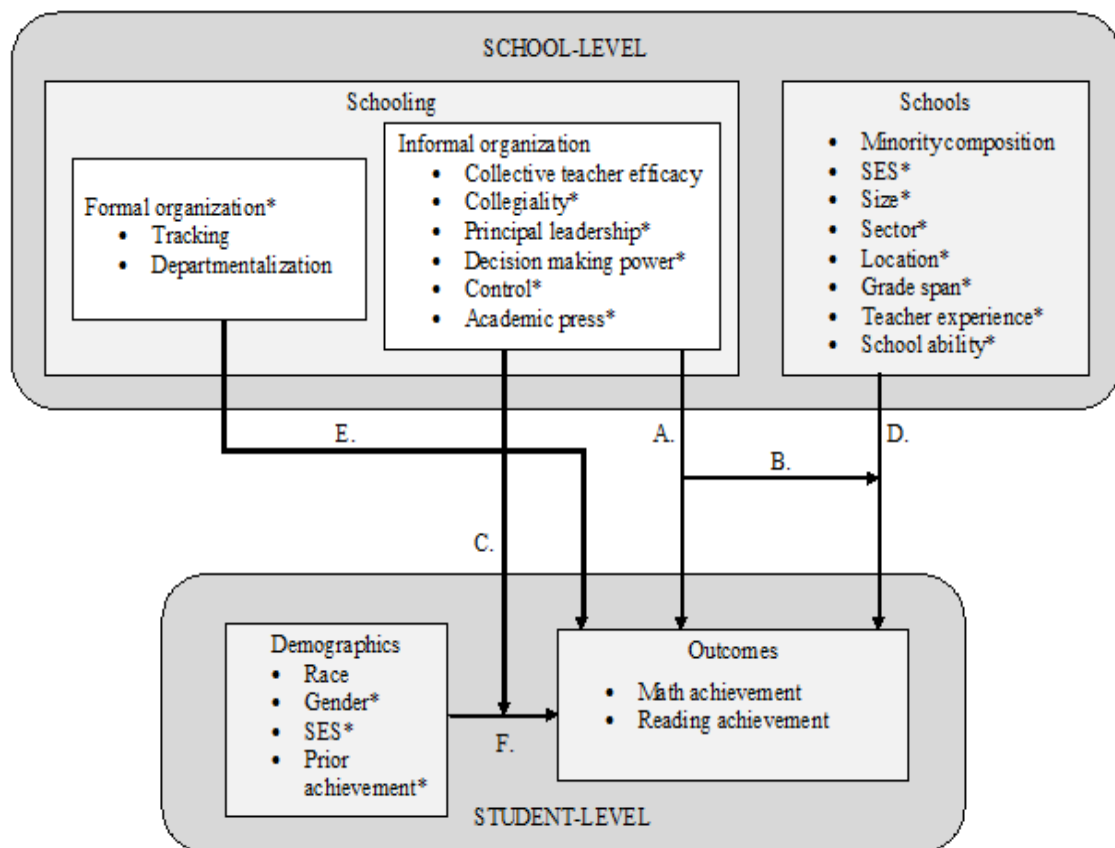
Importance of Study

Though traditionally a local issue, the federal government has played a role in America's schools most notably during the past 50 years. After the launch of Sputnik in 1957, fear of the United States falling behind the rest of the world drew national attention to the quality of the educational opportunities offered to students. Logically, any type of educational policy and reform would look towards teachers as an area of interest and the federal government is no different, as evidenced through their reports and policies. For example, the authors of *A Nation at Risk* suggested that the quality of education would improve through performance-based salaries and career ladders for teachers (U.S. Department of Education, 1983). During the Bush administration, the *No Child Left Behind Act* in 2002 attempted to increase student achievement by emphasizing teacher quality through increased teacher education and certification. President Obama's blueprint for reauthorizing the *Elementary and Secondary Education Act* (2010) and the *Race to the Top* (2009) program both emphasize the need for rigorous teacher evaluation systems and teacher performance pay. These federally endorsed reports, policies, and programs have in common an emphasis on individual teachers as a means for improvement.

The point here is not to describe the historical role of the federal government on the nation's schools; rather, that informal aspects of schooling, such as collective teacher efficacy, are missing features in education policy and reform. This study seeks to understand whether cultivating informal aspects of schooling that impact teachers

collectively as a school, might also facilitate effective policies and reforms. The hope is that even if the composition of a school is associated with negative outcomes for students, the internal workings of a school can improve a situation that is difficult to change. For example, many segregated and low-income schools suffer from high teacher attrition rates. Some studies, such as Liu (2007) found that increasing teachers' decision-making abilities within schools can help alleviate this problem, and in turn, improve student achievement in these populations.

Lastly, this study focuses on early high school achievement. Early high school is a significant time period because these years are crucial for school persistence and life outcomes. Students with poor academic performance are more likely to be held back a grade level, which is a strong predictor of a student dropping out of high school (Goldschmidt & Wang, 1999). Moreover, students who do drop out of high school leave early on in their high school careers (Somers & Piliawsky, 2004). Lastly, due to compulsory education laws that require young people to remain in school till early high school, this sample of students is less exceptional than a sample of students who have persisted till the end of high school.

Overview of Dissertation

The following chapters of this study include a review of the literature, a description of the methodology, the results of the analysis, and an interpretation of the results. Chapter 2 provides a literature review on prior research on collective teacher efficacy. This chapter narrows the extant literature to U.S. studies that treat collective teacher efficacy as an organizational property, are peer-reviewed, and use quantitative

methods to examine this construct. Chapter 3 describes the quantitative method of

analysis, the dataset used for this study, a description of the student sample and schools

included in this analysis, and the analytic strategy. Chapter 4 presents the results of the

analysis in regards to the research questions outlined in this chapter, and concludes with a

discussion of the results in Chapter 5.

CHAPTER 2: REVIEW OF THE LITERATURE

This chapter provides a review of the quantitative literature on collective teacher efficacy as an organizational construct. The literature for this review was gathered from a variety of databases: Education Research Complete, EBSCO, ERIC, EconLit, PsycINFO, and SocINDEX. I ran multiple searches with various combinations of the terms "collective," "teacher," and "efficacy." Other search terms included "sense of efficacy," "social organization," and "school organization." When limiting the search to peer-review articles, this search retrieved over 200 articles.

Collective teacher efficacy is not only examined in schools in the U.S., but this construct is widely examined abroad, as well. Studies in Turkey (e.g., Demir, 2008), Israel (e.g., Lev & Koslowsky, 2008), the United Kingdom (e.g., Parker, Hannah, & Topping, 2006), Canada (e.g., Ross & Gray, 2006; Ross, Hogaboam-Gray, & Gray, 2004), Norway (e.g., Skaalvik & Skaalvik, 2007; 2009), and Italy (e.g., Caprara, Barbaranelli, Steca, & Malone, 2006) demonstrate the recent interest in this field in other countries. These studies were excluded from this review because school systems vary greatly between countries and may not provide generalizable results for the educational system in the United States. Many countries abroad are significantly more homogeneous than in the states, organize students in different ways, and grant teachers different social status than schools in the United States. Furthermore, the U.S. has a unique history of inequality in the way our students are segregated by location, and in turn, their schools.

For these reasons, I excluded studies of collective teacher efficacy in countries outside of the United States from this analysis. When limiting my search to U.S. schools, the literature pool decreased to 29 articles.

Upon further examination, however, many of these articles were not studies on collective teacher efficacy, and were therefore excluded.[5] Studies on collective teacher efficacy that did not treat this construct as a school-level construct were excluded as well. For example, Zambo and Zambo (2008) examined 63 fourth through tenth-grade teachers to assess both their individual and collective efficacy. They described collective teacher efficacy as a "teacher's belief about his or her colleagues' effectiveness" (p. 160). Although they argued that this construct goes beyond the individual to focus on the faculty as a whole, it was not treated as a school trait. Rather, they focused on a teacher's individual opinion of their colleagues, and treated it as such when they did not aggregate this measure to the school level. Zambo and Zambo's study of collective teacher efficacy was therefore, not an examination of a school's organizational structure, but more an examination of the internal psychology of the individual teacher. Chambers and Cantrell (2008) also examined collective teacher efficacy after a professional development program. Similar to Zambo and Zambo, they conceptualized collective efficacy as an individual teacher property. Studies such as these were excluded from this literature review because these investigations into collective teacher efficacy were conceptually and methodologically different than investigations into collective teacher efficacy as a school-level property.

_____

[5] One example is Ware and Kitsantas (2011). Their measure of collective teacher efficacy was six items that assessed teacher perceptions of their decision making ability. Because Ware and Kitsantas' measure of collective teacher efficacy was not similar to the definition used here, I omitted this study from this review.

Lastly, I further limited my pool of literature by selecting studies that took place within the context of k-12 schools, pertained to general education (i.e., special education, physical education), and were quantitative in methodology. Ultimately, I was left with 12 studies for this literature review.[6] I utilized the snow-ball technique to ensure that no studies on collective teacher efficacy, as a school-level construct, were excluded.

This chapter is divided into seven sections. The first section provides a broad stroke of the literature I reviewed for this chapter. Next, I illustrate the multiple ways in which researchers have measured collective teacher efficacy. The third section contains research that examined factors that impact collective teacher efficacy within a school, followed by a section that describes how collective teacher efficacy affects student achievement. Fifth, I review studies that examine other outcomes, such as individual teacher efficacy and parental involvement. I then provide a small quantitative analysis using the current literature on the mediating effect of collective teacher efficacy. I conclude this chapter with the limitations of the literature on collective teacher efficacy and how this study contributes to the existing literature.

Broad Conceptualization of the Literature

Of these 12 studies on collective teacher efficacy, three were written in the late 1980s/early 1990s: Hoover-Dempsey, Bassler, & Brissie (1987), Newmann, Rutter & Smith (1989), and Lee, Dedrick, & Smith (1991). Starting in 2000, Roger Goddard and

---

[6] Some literature on professional learning communities (PLC) and teacher efficacy were initially examined for this literature review. Due to the parameters set on the literature search, this body of literature was excluded from this study. Many of the concepts of professional learning communities are very similar to the idea of informal organization of schooling, however.

colleagues from Ohio State University (William Hoy, Anita Woolfolk Hoy, and Scott Sweetland) began to extensively research this construct. While others had written studies on teacher efficacy during this ten year gap in the literature (from early 1990s to 2000), studies during this time period only examined efficacy as an individual, psychological aspect of schooling, as opposed to an organizational one (some examples include Coladarci, 1992; Moore & Esselman, 1992 & 1994; Pajares, 1996; Woolfolk & Hoy, 1990).

Prior literature includes examinations of factors that influence collective teacher efficacy, and the ways in which collective teacher efficacy influences other factors, including student achievement. Researchers used multiple quantitative methods to investigate this construct, including path analysis, multiple regression, and hierarchical linear modeling (HLM). Education researchers also used multiple methods to measure collective teacher efficacy; the next section describes the genesis and development of these measures.

Measures of Collective Teacher Efficacy

Multiple scales have been created to measure collective teacher efficacy. In general, all of these scales tap into teachers' feelings about either the efficacy of teachers in general, the ability of themselves personally, or the ability of the teachers in their schools to affect positive change in students. While these measures vary to some degree, Bandura (1993) suggests that all of these approaches are appropriate in evaluating collective efficacy and organizational performance, as long as these measures are aggregated to the school level.

The most common collective teacher efficacy scales were initially based on Gibson and Dembo's (1984) individual teacher efficacy scale. Their initial pilot study using the teacher efficacy scale consisted of 53 items and was administered to 90 teachers. Through a factor analysis, the scale was reduced to a 30-item questionnaire in a Likert format (1 = strongly disagree to 6 = strongly agree). Gibson and Dembo administered the 30-item individual teacher efficacy scale to 208 teachers selected from 13 elementary schools within two neighboring unified school districts. With these results, the authors employed another factor analysis to analyze the underlying structure of the teacher responses. Two factors were extracted from these data. The first, personal teaching efficacy, consisted of nine items and represented a teacher's belief that he/she had the skills and abilities to bring about student learning. One sample item of this factor included "When I really try, I can get through to most difficult students" (p. 573). The second factor, general teaching efficacy, consisted of seven items and represented a teacher's belief about the general relationships between teaching and learning. One sample item of this factor included "Even a teacher with good teaching abilities may not reach many students" (p. 573). These two factors were consistent, most notably, with the Rand study (Armor, et al., 1976) on teacher efficacy, which separated personal teaching efficacy from general teacher efficacy. Gibson and Dembo found that the teacher efficacy scale could be separated into two parts or combined together. They created a more parsimonious model and pared down their scale to 16 items. A measure of internal consistency generated Cronbach's alpha coefficients[7] of 0.79 for all 16 items, 0.78 for personal teaching efficacy alone, and 0.75 for general teaching efficacy alone.

---

[7] Cronbach's α is a commonly used measure of internal consistency, which refers to how

Goddard and his colleagues subsequently adapted Gibson and Dembo's scale to measure collective teacher efficacy. Goddard, Hoy, and Hoy (2000) suggested that two elements were key in developing collective teacher efficacy: assessment of group competence, and analysis of teaching task. These authors postulated that teachers analyze the competence of their colleagues' skills, methods, training, and expertise. A teacher's assessment of the group's competence interacts with a teacher's analysis of what constitutes success in their school, the limitations that must be overcome, and what resources are available to succeed. Goddard and his colleagues also cited previous studies that suggested that teachers may express different efficacy beliefs depending on whether the outcomes were described as positive or negative. Starting with Gibson and Dembo's 16-item scale, Goddard et al. identified four different categories of questions: group competence worded positively, group competence worded negatively, teaching task worded positively, and teaching task worded negatively. These authors also changed the individually worded items to a group orientation. For example, "I can reach a difficult student" was altered to assess collective teacher efficacy when it was changed to "Teachers in this school can reach a difficult student" (p. 487). When applying these categories to Gibson and Dembo's scale, they found that only positively worded group competence and negatively worded teaching task groups were represented. Therefore, Goddard and his colleagues generated items to fulfill all four categories, albeit unequally, in order to provide a more balanced collective teacher efficacy scale. Goddard et al. field tested their collective teacher efficacy scale on a sample of 70 teachers, one from each of

well the items of an instrument fit together (Pett, Lackey, & Sullivan, 2003). Nunnally (1978, as cited in Santos, 1999) stated that a reliability coefficient of 0.7 or higher was considered to be acceptable.

70 schools in five states. After they employed a factor analysis, omitted redundant items, added additional group competence items, and added items that gauged the difficulties and resources for teaching, these authors created a 21-item scale named the Collective Teacher Efficacy scale.

Recognizing that the Collective Teacher Efficacy scale did not maintain equal coverage of the four different categories mentioned above, Goddard (2002) created a more parsimonious version of the 21-item scale. He selected the items with the largest loadings for each of the four efficacy categories based on a principal axis factor analysis. A one-factor solution was extracted which contained 12 items. Goddard found that the abbreviated version was highly correlated to the original 21-item scale (r = 0.983), which suggested that the omission of 43 percent of the items would result in little, if any, change in meaning or reliability. Indeed, the internal consistency of the original versus the short form was quite similar (α = 0.96 and α = 0.94, respectively).

The majority of the literature on collective teacher efficacy used some form of the Collective Teacher Efficacy scale. Out of 12 studies, five studies used the short version, while three used the longer version. The use of this scale is not surprising because the majority of the work written on collective teacher efficacy has been written by the same few researchers (e.g., Goddard, R. Hoy, and A.W. Hoy).

Of the remaining four studies, two studies used their own scales. Tschannen-Moran and Barr (2004) developed their own collective teacher efficacy scale due to concerns that Goddard et al.'s Collective Teacher Efficacy scale artificially drove down scores of schools in more challenging environments. Tschannen-Moran and Barr's Collective Teacher Belief Scale contained two subscales: instructional strategies and

student discipline. Teachers were individually asked about their perceptions of collective efficacy rather than their personal efficacy beliefs. Instructional strategy questions included questions such as "How much can teachers in your school do to produce meaningful student learning" (p. 198)? Student discipline questions included questions such as "To what extent can school personnel in your school establish rules and procedures that facilitate learning" (p. 199)? Their 12-item scale had a reliability of 0.97.

Hoover-Dempsey, Bassler, and Brissie (1987) also created their own scale to measure collective teacher efficacy. Their 11-item measure was part of a larger 164-item questionnaire. One sample item was "I feel that I am making a significant difference in the lives of my students" (p. 425). Their internal consistency, however, was the second-lowest of all the collective teacher efficacy scales with an alpha of 0.87. Although still a reasonable level of reliability for a psychometric scale, the Hoover-Dempsey, Bassler, and Brissie scale demonstrate that differences in findings between studies may be due, at least partially, to differences in the psychometric properties of scales.

Two studies, Newmann, Rutter, and Smith (1989) and Lee, Dedrick, and Smith (1991) used the same questions to create their collective teacher efficacy measure; both were based on the same national dataset, the 1980 High School & Beyond study. Their measure of collective teacher efficacy consisted of four questions with the lowest internal consistency ($\alpha = 0.73$) of all the scales. Lower levels of reliability are not uncommon when scales are developed from general-purpose surveys, such as HS&B and subsequent national surveys of education. The HS&B items asked teachers about their personal efficacy including questions such as "How successful do you feel in educating students?" and "I look forward to working every day" (p. 205). These researchers examined

collective teacher efficacy by aggregating teacher responses to these individualized

questions.

Summary

The following studies show that early researchers of collective teacher efficacy

based their measurements on existing individual teacher efficacy research. While many

researchers adapted these scales over time, others constructed their own collective teacher

efficacy scales and measures based on their conceptualizations of this construct. The

internal consistencies of the measurements described above ranged from as high as 0.96

to as low as 0.73, which is evidence that collective teacher efficacy can be measured

using survey results, though with varying degrees of reliability. Additionally, the items

used in these measures are similar to the survey items that I intend to use in my study.

Factors that Influence Collective Teacher Efficacy

Many school factors contribute to a school's collective sense of teacher efficacy.

Researchers found that certain school demographics, structure, and formal and informal

organization of schooling have had some level of impact on collective teacher efficacy.

One of the first studies written on the impact of school organizational features on

collective teacher efficacy used the national dataset, High School and Beyond (HS&B).

Newmann, Rutter, and Smith (1989) worked with data from the 1984 Administrator and

Teacher Survey, which was a subset of teachers and schools from HS&B. Their national

sample included up to 30 randomly selected teachers from 353 public high schools. In

creating their collective teacher efficacy outcome measure, they employed a factor

analysis by creating a composite score at the individual (teacher) level, and then

aggregated the scores to the school level. A sample item from this questionnaire includes "I sometimes feel it is a waste of time to try to do my best as a teacher" (p. 228). The authors included five school demographic variables and 10 organizational variables into their models. School demographic variables included student ability (prior achievement), school size, percentage of low socioeconomic status (SES) students, percentage of White students in a school, and urbanicity. Organizational variables included level of student misbehavior and school order, administrative responsiveness and support, teacher's feelings of their level of influence in the school, encouragement of innovation within the school, whether knowledge is shared amongst teachers, whether teachers helped one another, feelings about their principal's leadership, types of in-service programs, amount of meeting time with other teachers, and staff development opportunities. In addition to these variables, they included the within-school variance of teacher attitudes. This within-school variance measure represented the level of consensus within a school regarding collective teacher efficacy. Due to a listwise procedure of deleting cases with missing variables, Newmann and his colleagues' final sample consisted of 288 schools.

Through multivariate regression where only school demographic variables were entered into the model, student average ability seemed to have the largest effect (0.40 SD) [8]; however, when school organizational variables were included in the analysis, school demographic variables had less predictive power. In their full model, student ability (0.23 SD), school order (0.17 SD), encouragement of innovation (0.19 SD),

---

[8] These measures were standardized, meaning that these measures were converted into Z-scores and recalibrated into a measure with a mean of 0 and a standard deviation (SD) of 1. One SD above and below the mean accounts for 68.2% of the sample. Measures in SD can also be interpreted as an effect size where 0.2 is considered to be a small effect, 0.5 is a medium effect, and 0.8 is a large effect (Cohen, 1988).

sharing of teacher knowledge (0.11 SD), and the absence of school consensus (-0.34 SD) had statistically significant effects on teacher efficacy. The authors speculated that the negative coefficient associated with the variability in school consensus was due to disparities and divisiveness that may be created when teachers within a school differed in their sense of efficacy. Newmann and his colleagues also found that as the percentage of White students in the school increased, collective teacher efficacy was predicted to decrease (-0.17 SD). The authors suspected that, on average, teachers in schools with high minority populations may make special efforts that increase teachers' sense of collective efficacy.

Hoy, Sweetland, and Smith (2002) expanded upon Newmann et al.'s findings with their examination of the relationships between academic press, collective teacher efficacy, school SES, and mathematics achievement. These authors defined academic press of a school as the extent to which the school was driven by a quest for academic excellence. They hypothesized that school SES and academic press in a school had an independent relationship to collective teacher efficacy. Through a path analysis, they found that both academic press and SES had significant and direct effects on collective teacher efficacy; the effect of academic press was stronger, however, than the effect of SES (0.56 SD, $p < 0.01$ compared to 0.25 SD, $p < 0.01$).[9]

Goddard, LoGerfo, and Hoy (2004) also examined the effect of school context on student achievement and possible mediating effects of collective teacher efficacy. The exploration of these mediating effects required the researchers to also examine the

---

[9] Although the primary dependent variable for this study was student achievement, collective teacher efficacy was also used as a dependent variable as part of this path analysis. I used the results of the path analysis to describe factors that influence teachers' sense of collective efficacy. This is also true of several of the studies that follow.

relationship between different school context variables and collective teacher efficacy. Similarly to Hoy et al. (2002), these authors used a path analysis to analyze data drawn from students and teachers in 96 high schools in a large Midwestern state; however, instead of only examining mathematics scores, these authors extended their analysis to include other subjects. They focused on two identical models with different twelfth-grade outcome measures. One model focused on verbal achievement, which included the proportion of students in a school that passed the reading, writing, and social studies assessments. The other model focused on mathematics achievement, which included the proportion of students in a school that passed the twelfth-grade mathematics and science assessments.

Goddard, LoGerfo, and Hoy included five school context variables in their analysis. In addition to collective teacher efficacy, these school variables were school minority enrollment, urbanicity, SES, size, and prior achievement. These authors measured prior achievement as the proportion of students in the school who passed the ninth-grade subject assessments. In their verbal assessment model, they found that school SES ($\beta = 0.35$, $p < 0.05$), school size ($\beta = -0.12$, $p < .05$) and prior achievement ($\beta = 0.44$, $p < 0.05$) were significant predictors of collective teacher efficacy; they also found similar results for their mathematics and science model. The proportion of minority students in the school and school urbanicity had no statistical significance.

While Goddard et al. examined high schools within a Midwestern state, Goddard (2001) examined elementary schools in one Midwestern district. He used several means-as-outcomes models – multilevel models that include only school-level characteristics – to test the effect of school-level variables as predictors of between-school variability in

collective teacher efficacy. The proportion of African American students and low SES students among schools was negative and significantly related to collective teacher efficacy when in the models alone. However, when both were included in a model with prior achievement (measured as third-grade district test scores), prior achievement was the only significant predictor among schools of collective teacher efficacy.

In another study focused on prior-year student achievement, Cybulski, Hoy, and Sweetland (2005) combined economic theory with organizational theory in their examination of 146 elementary schools in Ohio. While not a random sample, they studied a relatively representative sample of rural, suburban, and urban schools. Their economic variables of interests were student instructional and student services ratios. The student instructional ratio was the proportion of student costs in the classroom compared to administrative and operational costs, whereas the student services ratio was the proportion of money for instructional activities both inside and outside the classroom compared to other administrative and operational costs. These authors, through a path analysis, aimed to examine the relationship between student instructional ratio, student services ratio, school SES (as measured as the proportion of students not receiving free and reduced lunch), collective teacher efficacy, and student achievement. Their student achievement measures were fourth-grade reading and mathematics proficiency exams for the year preceding the study and the year of the study.

These authors examined four path models based on different combinations of the two economic measures and the two achievement measures. Consistent with past studies, Cybulski et al. found prior achievement to have a statistically significant effect on collective teacher efficacy for all model combinations (ranging from $\beta = 0.15$ to 0.16).

However, the effect of SES was more than four times greater than the impact of prior achievement ($\beta = 0.70$, $p < 0.05$ for all four models). The economic measures had no statistical effect on collective teacher efficacy.

Adams and Forsyth (2006) also examined collective teacher efficacy, prior achievement, and SES, with the addition of school level and school structure. Their sample included 22 elementary schools, 30 middle schools, and 27 high schools in one Midwestern state. Ten teachers were randomly sampled from each school to participate in their survey, which yielded a 69 percent teacher-return rate. As for their variables of interest, the authors used the proportion of students on free and reduced lunch as a proxy measure of school SES, and their prior achievement scores were based on a school's academic performance index. Ninety percent of the academic performance index was based on the state mandated criterion reference test, and 10 percent was based on student attendance. Their school structure variable measured whether teachers perceived that the bureaucratic structures in their school were enabling or hindering. Enabling structures have rules, regulations, and procedures that are "helpful and lead to problem solving among members" (Hoy & Sweetland, 2000, p. 531, as cited in Adams & Forsyth, 2006, p.634). This measure consisted of a 12-item survey that included responses to questions such as "administrative rules in this school are guides to solutions rather than rigid procedure" (p. 634).

Through multivariate regression, these authors found that prior academic performance (0.46 SD) and an enabling school structure (0.36 SD) had positive effects on collective teacher efficacy, while the percentage of free and reduced lunch students in a school had a negative effect (-0.23 SD). Moreover, high school teachers were found to

have lower feelings of collective efficacy compared to teachers in lower-grade levels. Prior achievement accounted for 54 percent of the variance in collective teacher efficacy, while the other school factors accounted for 20 percent, combined. Their research is consistent with others in highlighting the importance of prior achievement and informal organizational factors in influencing feelings of collective efficacy.

The majority of the studies described above attempt to uncover the ways in which student composition, school structure, and organizational aspects of the school influence collective teacher efficacy. Goddard and Skrla's (2006) research is unique in that this study examined the effect of teacher demographics. Their study was based in a diverse urban school district in the southwestern part of the country. They examined 1,981 teachers in 41 kindergarten through eighth-grade schools. Students in these schools, on average, were 53 percent Latino/a and 35 percent Black. Teacher characteristics included their race and ethnicity, gender, and their years of experience. School level variables included the proportion of socioeconomically disadvantaged students, the proportion of Latino/a and Black students, prior-year's level of student academic proficiency in the state reading assessment, and the proportion of students enrolled in gifted programs.

Because Goddard and Skrla aimed to examine teachers within schools, they employed hierarchical linear modeling. In their within-school model, they found that Black and Latino/a teachers, compared to nonminority teachers, had more positive beliefs about collective teacher efficacy in a school. Teachers with 10 or more years of experience also had a positive effect on collective teacher efficacy, compared to teachers at midcareer. While these results were noteworthy, these teacher-level variables only explained four percent of the within school variance. Their full model retained the same

results of the within school model, but also found that the proportion of Latino/a teachers, the proportion of gifted students, and prior reading proficiency were positive predictors of differences among schools in their collective teacher efficacy.

Lee, Dedrick, and Smith (1991) also examined the relationship between individual teacher variables and their school context. While these authors used the same dataset as Newmann et al., their study differed in focus and methodology since Lee et al. included a distinction between public and Catholic high schools. Moreover, Newmann et al. examined collective teacher efficacy using multivariate analysis, whereas Lee et al. used hierarchical linear modeling to account for the nested nature of school data.

Lee and her colleagues' study involved 8,488 full-time teachers in 354 high schools. Of these schools, 307 were public and 47 were Catholic high schools. These authors used identical items as Newmann et al.'s collective teacher efficacy construct.[10] In addition to school sector, they included organizational variables such as school SES, school size, teachers' sense of control, sense of disorder, and sense of community in the school. In order to examine whether the effect of teacher control on efficacy varied by school characteristics, they interacted these school variables with the teacher control slope.

In their unconditional model, they found that teacher control and student ability were both strongly related to within-school beliefs about teacher efficacy. With the addition of sector in their first between-school model, Catholic schools were found to

---

[10] Although Lee et al. and Newmann et al. used the same items to construct collective teacher efficacy, Lee et al. initially intended to use these items to study both teacher efficacy and job satisfaction. Factor analysis rotated the items into one factor, thus combining the individual indicators into a single construct. For consistency, Lee et al. named the factor teacher efficacy, although they "admit that the labeling is somewhat arbitrary" (p. 195).

have higher collective efficacy (0.29 SD) than public schools. However, the effect of teaching in a Catholic school became nonsignificant once other school organizational variables were included in the model.

In their final school organization model, Lee et al. found that teachers in high SES schools, large schools, and schools with a strong sense of teacher control, strong feelings of community, strong sense of principal leadership, and low feelings of disorder felt more efficacious. Though the effect sizes were small (the largest being 0.09 SD), their research showed that the difference in collective teacher efficacy between Catholic and public schools could be explained by both formal and informal organizational differences. The only school measure that that was significantly related to the teacher control slope was principal leadership. In other words, in schools with strong principal leadership, the relationship between teacher's feelings of control and teacher efficacy is even stronger. Unlike Goddard and Skrla, these authors omitted teacher and school demographic measures from their final models because they found no relationship to teacher efficacy or control.

Summary

Across these studies, I found consistent evidence that school prior achievement had a large effect on the collective efficacy of teachers in a school. Certain informal organizational aspects of schooling, such as sense of order in the school, were also found to positively impact collective teacher efficacy. The effect of a school's socioeconomic status and minority composition was mixed: while some studies found statistical significance in these variables, others found that these variables had no impact on collective teacher efficacy.

Collective Teacher Efficacy and Student Achievement

While the link between individual teacher efficacy and achievement is well established, there has been less research on the link between collective teacher efficacy and student achievement (Tschannen-Moran & Barr, 2004). Through my literature search, I found only five studies that quantitatively examined this relationship. Tschannen-Moran and Barr's (2004) study of 66 middle schools in Virginia was the least sophisticated in statistical methodology. With eighth-grade mathematics, writing, and English state standardized test scores as the outcomes, they ran three regression models each with only collective teacher efficacy and school SES as their independent variables of interest. Their SES measure represented the proportion of students in a school that received free and reduced price lunch. These authors found that SES had a large negative relationship to all of the subject scores. Collective teacher efficacy was only significant in the writing assessment when controlling for SES (0.273 SD, $p < 0.001$), and accounted for 28 percent of the variance on the writing test.

In a more statistically sophisticated examination of collective teacher efficacy and achievement, Goddard, Hoy, and Hoy (2000) aimed to find whether collective teacher efficacy explained differences in student achievement between schools. Their sample consisted of 7,016 students of 452 teachers within 47 elementary schools in one large, urban Midwestern district. Student achievement was measured by the state mathematics and reading standardized test. Other student variables included gender, race/ethnicity, and free or reduced priced lunch status. Employing hierarchical linear modeling, these authors found that collective teacher efficacy was positively related to both mathematics and reading achievement ($b = 8.62$, $p < 0.001$ and $b = 8.49$, $p < 0.001$, respectively).

Low-income students and Black students were predicted to have lower achievement scores in both subjects. Lastly, females were predicted to score higher than males on the verbal assessment, but not on the mathematics test.

Goddard's 2001 study of urban elementary school children in a Midwest state ran similar models as Goddard, Hoy, and Hoy (2000) but also controlled for prior achievement. Goddard found that collective teacher efficacy still had a positive impact on mathematics and reading achievement. In addition, females were predicted to score higher than males, and African American students were predicted to score lower than non-African American students on both subject tests.

Goddard extended his study by also predicting the impact of efficacy consensus within a school (omitting collective teacher efficacy) with student achievement. He defined efficacy consensus as the level of variation in teacher responses within the school. When he replaced efficacy consensus with collective teacher efficacy in the same models, consensus was not statistically significant, while all other variables remained significant and were consistent with the collective teacher efficacy models.

Similar to previous studies, Goddard, LoGerfo, and Hoy's 2004 study also found collective teacher efficacy to impact student achievement. Their examination of 96 high schools in a large, Midwestern state found, through regression analysis, that a one unit increase in collective teacher efficacy would increase 12th grade mathematics and verbal achievement by 0.23 and 0.24 SD, respectively.

Likewise, Hoy, Sweetland, and Smith (2002) found similar results in their investigation of 97 high schools in Ohio. After controlling for SES and academic press, these authors found that collective teacher efficacy had an effect on school mathematics

achievement ($\beta = 0.51$, $p < 0.01$). This effect was twice as strong as the effect of SES ($\beta = 0.21$, $p < 0.05$).

Summary

Since collective teacher efficacy and student achievement deals with schools and their students, investigations into this phenomenon require statistical methods that can account for the multilevel nature of these data. Statistically speaking, in order to avoid aggregation bias, misestimated standard errors, and heterogeneity of regression issues that may occur when individual-level characteristics are aggregated to the group level, studies involving nested data should employ hierarchical linear modeling (Goddard & Goddard, 2001). Out of the five studies that dealt with student achievement, only Goddard, Hoy, and Hoy (2000) and Goddard (2001) employed hierarchical linear modeling to examine students within schools. To their credit, for some of the earlier studies, multi-level modeling was not widely used as a method for accounting for these differences.

Though the literature on collective teacher efficacy and student achievement is small, these researchers consistently found a positive and statistically significant relationship between these two variables, even after controlling for a number of school factors, such as school SES, student minority composition, school's prior academic achievement, and academic press.

Collective Teacher Efficacy and Other Outcomes

Not only is collective teacher efficacy known to impact student achievement, but some studies found this construct to impact other non-student outcomes, such as individual teacher efficacy and parental involvement.

Goddard and Goddard (2001) examined the relationship between collective teacher efficacy with individual teacher efficacy. The authors gathered these data from a survey of teachers in a large urban school district in the Midwest. Their final sample involved 438 teachers in 47 elementary schools of varying grade levels. In order to assure anonymity, these researchers did not gather teacher demographic information but did gather school-level contextual variables. These variables included the proportion of students receiving free and reduce lunch, proportion of minority students in a school, prior achievement (measured as 3rd grade students' mathematics score on the state assessment taken one year prior to the study), and school size. These authors used a five-item scale based on Gibson and Dembo (1984) as their measure of individual teacher efficacy, while they used a 21-item collective teacher efficacy scale by Goddard, Hoy, and Hoy (2000).[11]

Using HLM as their methodology and individual teacher efficacy as the outcome, these authors tested each school-level characteristic separately and then tested a model that combined these characteristics. Goddard and Goddard found that individually, the proportion of low-income students (-0.10 SD), prior achievement (0.11 SD), and collective teacher efficacy (0.19 SD) were related to individual teacher efficacy. In a

---

[11] As described in a prior section, Goddard, Hoy, and Hoy's (2000) 21-item scale was based on Gibson and Dembo's (1984) scale; thus, the use of collective teacher efficacy in predicting individual teacher efficacy in this study may be positively biased.

combined model of these three significant variables, collective teacher efficacy was the only significant predictor of individual teacher efficacy (0.25 SD). These authors concluded that the variation between schools in individual efficacy may be explained by collective teacher efficacy because individual teacher efficacy was higher in schools where collective teacher efficacy was also higher.

Not only did collective teacher efficacy affect teachers individually, but has also been found to affect parental involvement. Hoover-Dempsey, Bassler, and Brissie's (1987) examination of efficacy and parent involvement drew their data from teachers and principals in 66 elementary schools from a large mid-Southern state in the U.S. Using multivariate regression, these authors examined school demographics (such as school SES), class size, and organizational factors (such as teacher efficacy and principal's feelings about their teachers' efficacy) to explain five parent involvement variables. These five dependent measures included the average number of students whose parents attended a parent-teacher conference in the school; the average number of parent volunteers; the average number of students whose parents spent time on school related tasks at home; the average number of students whose parents provided home instruction on a plan devised by the teacher; and the teacher's response to the item, "most of my students' parents support the things I do." Collective teacher efficacy had a strong positive effect on parent involvement, specifically for attendance at parent/teacher conferences ($\beta = 0.355$, $p < 0.001$), parent volunteering ($\beta = 0.322$, $p < 0.01$), parent tutoring at home ($\beta = 0.344$, $p < 0.01$), and teacher's feelings of support from parents ($\beta = 0.552$, $p < 0.001$). School SES also had an impact for parent/teacher conferences, parent volunteers, and teacher's feelings of support from parents. These authors suggested that

high efficacy may demonstrate a sense of professionalism and security in the teaching role. Such confidence would enhance a teacher's ability to discuss their teaching program and goals at conferences. In addition, when these teachers ask parents for help, others may interpret this outreach as an accompaniment to teaching, and not as a sign of teaching inadequacy. Lastly, they suggested that teachers with high efficacy may minimize the potential perceptions of creating a threatening environment for parents, which aids in the parent-teacher relationship.

Summary

While the studies in this section are limited, I found that collective teacher efficacy as an informal organizational property impacts schools by increasing individual teacher's feelings of their efficacy, and by increasing parent involvement. Schools with strong collective teacher efficacy seem to affect individual teacher's behaviors and attitudes to make schools more inviting for parents to partake in their child's schooling. Or, conversely, it may be that parent involvement and individual feelings of efficacy make it easier for teachers to build a collective sense of efficacy.

Collective Teacher Efficacy as a Mediator

Investigations into school characteristics, such as collective teacher efficacy, and the ways that it can indirectly lessen the impact of racially and economically segregated schools could be important. Many studies have shown that students fare worse academically in high minority enrollment and low income schools (some examples include Bankston and Caldas 1996 and 1998, Hanushek and Rivkin 2009, and Gamoran 1987). Unfortunately, none of the studies selected for this literature review formally test

for the mediating or moderating effect that collective teacher efficacy may have on these types of schools. With information provided by three studies that use path analysis, however, I am able to approximate the indirect role that collective teacher efficacy may play in equalizing unequal aspects of schools. For this section, I use Kenny's (2011) method of calculating the indirect and total effects of a mediator model, which is outlined in Appendix A.

Goddard, LoGerfo, and Hoy (2004) used a path analysis in their examination of collective teacher efficacy and high school student achievement in one Midwestern state. While the indirect effect of school socioeconomic status and school verbal achievement through collective teacher efficacy seemed relatively small (0.08 SD), collective teacher efficacy mediated 26 percent of the total effect of SES on twelfth-grade verbal achievement. Similarly, for math and science achievement, the indirect effect of collective teacher efficacy was 0.06 SD, but the proportion of the total effect of SES that was mediated by collective teacher efficacy was 25 percent. These authors also examined the relationship between high minority population schools and achievement, but they found no direct effects between these two variables, and no indirect effects with collective teacher efficacy.

Cybulski, Hoy, and Sweetland (2005) also examined the mediating effect of collective teacher efficacy on the relationship between school SES and two student achievement outcomes, mathematics and reading. With data based on their sample of elementary schools in Ohio, I found that the indirect effect of collective teacher efficacy ranged from 0.15 SD to 0.19 SD. The proportion of the effect ranged from 35 to 33.2

percent, respectively, which was slightly higher than Goddard, LoGerfo, and Hoy's high school study.

Lastly, in their examination of 97 high schools in Ohio, Hoy, Sweetland, and Smith (2002) examined collective teacher efficacy as a mediator using path analysis. This study suggests an even larger mediating effect of collective teacher efficacy than the aforementioned studies. Based on the results from their study, I found that 37.8 percent of the effect of SES on mathematics achievement was mediated by collective teacher efficacy.

Summary

Using calculations based on Kenny (2011), I found that collective teacher efficacy had a significant mediating effect on the effects of socioeconomic status of schools and students' achievement, at both the elementary and high school levels. I also discovered that these studies either lacked information by which to estimate indirect effects or found no mediating effects of collective efficacy for high minority enrollments. Although not conclusive, these findings suggest that teacher collective efficacy may mediate, that is, help to minimize the effects of the socioeconomic status on student achievement, and warrants further investigation.

Chapter Summary and Limitations of Prior Research

The existing literature on collective teacher efficacy as an organizational property of the school demonstrates the important role that school characteristics play in student learning. Past research described collective teacher efficacy as an independent and positive component to improving student achievement. The literature review revealed

that many school characteristics, such as prior student achievement, SES, and teacher's sense of collegiality, also impacted teachers' feelings of their collective efficacy. These results were consistent regardless of the scale used to measure collective teacher efficacy, and the location of the populations of interest. With information garnered from additional analyses of the literature, I found that collective teacher efficacy had a significant and potentially meaningful mediating effect on the relationship between school socioeconomic status and student achievement.

The literature on collective teacher efficacy is dominated by a few researchers, namely Goddard, W. Hoy and A.W. Hoy. These three researchers alone have published over half of the quantitative examinations of collective teacher efficacy as a school-level construct. While they greatly increased the knowledge base of collective teacher efficacy, many aspects about its impact on students are still unknown. For example, I found no investigations of collective teacher efficacy's direct effect on student achievement with the use of a national dataset – only examinations of schools within districts or states. Moreover, while collective teacher efficacy's role as a mediating variable was confirmed, its role as a moderator has yet to be examined, especially in regards to minority students and schools. This study fills these two gaps in the literature on collective teacher efficacy to gain additional insight into this construct on a national sample of students. In the following chapter, I describe the sample of students and schools in more depth, and provide the quantitative framework in which I answer the research questions presented previously.

CHAPTER 3: METHODOLOGY

The purpose of this chapter is to describe the data and methods used to analyze the effect of collective teacher efficacy, both as a moderator and as an independent variable, on high school student achievement. As mentioned in Chapter 1, my research questions are as follows:

1. Does collective teacher efficacy have a positive impact on student achievement?

2. Does a gap exist between minority students and majority students in regards to achievement within schools? If so, do high schools with high collective teacher efficacy have an equalizing effect on this achievement gap within schools?

3. Do high schools with large minority concentrations have different achievement levels compared to schools with low minority concentrations? If so, do high schools with high collective teacher efficacy have an interacting or moderating effect on the relationship between schools with high minority populations and student achievement?

In order to answer these research questions, I used the National Education Longitudinal Study of 1988 base year and first follow-up questionnaires.

National Education Longitudinal Study of 1988

The National Education Longitudinal Study of 1988 (NELS:88) intended to provide trend data of students' transition between middle school to high school, and their

subsequent entrance into the workforce or postsecondary education (Department of Education, n.d.). The first year of the study, the base year, took place in the spring of 1988. The base year sample consisted of 1,052 randomly selected public and private schools and 24,599 randomly selected eighth graders who attended these schools. A subset of these students was resurveyed in the subsequent waves of the study in 1990, 1992, 1994, and 2000. In addition, teachers, parents, and school administrators were also surveyed during these follow-up years. NELS:88 can be analyzed as a cross-section of students at one point in time, or longitudinally as a panel dataset (Curtin et al., 2002). For this study in particular, I used NELS:88 as a longitudinal dataset since I examined the same students during two time points. I drew data from the student base year and first follow-up questionnaires, the 1990 teacher questionnaire, and the 1990 school administrator questionnaire.

During the base year in 1988, approximately 24 eighth-grade students from each school were randomly selected, with the addition of two or three Asian and Hispanic students from each school. Since approximately 90 percent of eighth-graders moved to a new school during their transition into high school, students were traced to their new schools for the first follow-up. Students who dropped out of school were also asked to continue to participate in the follow-ups. The first follow-up student sample was also freshened[12] in order to be nationally representative of high school sophomores in 1990 (Curtin et al., 2002).

The first follow-up teacher survey was intended to inform and explain student behaviors, provide evaluations of their students' performance, provide teacher

---

[12] Additional students were added, or "freshened," to the base year sample of students to achieve a representative sample of tenth-graders in 1990 (Curtin et al., 2002).

demographic information, and provide teachers' perceptions of school climate and culture. Teachers were given approximately two weeks to complete a self-administered survey that was mailed to each teacher individually. This sample of teachers was not nationally representative of tenth-grade teachers across the country because participation was based on their students' participation during the base year in 1988. For each student involved in the study, two core subject area teachers were sampled. These subject areas were mathematics, English, science, and history (Ingels et al., 1992b).

School administrators were asked in the spring of 1990 to complete a self-administered questionnaire on the educational settings of the student participants. These administrators were given approximately two weeks to complete this survey. Although 23 percent of administrators failed to return the questionnaire in the spring, NELS:88 resent the survey in November 1990, which raised the return rate by an additional four percent. Interviewers contacted the remaining schools by phone to complete an abbreviated questionnaire. The final response rate for principals was 97 percent (Ingels et al., 1992a).

I selected NELS:88 as this study's data source due to the variables and data included in the survey. This dataset not only sampled a nationally representative group of students, but the survey followed many of these students over time. Furthermore, NELS:88 included information on these students' parents, teachers, and school administrators, which provides a rich picture of the conditions, background, and character of the student respondents and the schools that they attended. NELS:88 is also an ideal dataset because the creators of this survey composed six questions that aim to capture the concept of teacher efficacy (see Ingles et al., 1994b). Previous researchers have used

other national datasets to measure collective teacher efficacy, such as the High School &

Beyond Survey or the Schools and Staffing Survey (see Newmann et al., 1989 and Lee et

al., 1991); however, these surveys do not directly assess the potential role of collective

teacher efficacy on student content knowledge of core subjects. The use of NELS:88 also

allows me to utilize eighth-grade test scores as a base line in order to examine academic

growth from the beginning of high school till their tenth-grade year. Lastly, due to

compulsory education laws which require many students to remain in schools till the age

of 16, the NELS:88 first follow-up survey is appropriate for this study.

Missing Data

The analytic sample for this study only included students who were surveyed in

both the base year and first follow-up year, and who had tenth-grade school information.

In addition, I restricted the sample to schools that have at least five surveyed students, in

order to have reasonably reliable within-school sample sizes (Lee & Smith, 1995). Under

these parameters, a sample of 13,739 students remained from the original base year

sample size of 24,599; however, approximately 10.28 percent of the values in this dataset

were missing, and over 60 percent of students had at least one missing item. If using

listwise deletion methods to handle this missing data, I would drop all students and

schools with any missing data and I would be left with less than 40 percent of eligible

students for this study. Table 3.1 provides an analysis of the missing data, which

compares the full population of eligible students and schools in NELS:88 (Column 1), the

sample of students that would have been analyzed had I used listwise deletion (Column

2), and the sample of students that I would have dropped in the analysis due to missing

data (Column 3).

Table 3.1: Missing data analysis

Variable	1 Original Population M [a]	2 Potential Analytic Sample M [a]	3 Dropped Cases M [a, b]
	n = 13739	n = 5737	n = 8002
Student Variables			
10th Grade Reading	30.447	30.809	30.180 ***
10th Grade Math	43.547	44.165	43.072 ***
8th Grade Reading	27.057	27.429	26.793 ***
8th Grade Math	36.260	36.736	35.921 ***
Socioeconomic Status	-0.076	-0.033	-0.105 ***
	% [a]	% [a]	% [a]
Majority Student	79.7	82.3	77.9 ***
Minority Student	20.3	17.8	22.1 ***
Male	50.1	49.0	50.9 *
School Variables			
Public	94.9	93.7	95.7 ***
Urban	20.9	18.9	22.3 ***
Rural	36.9	39.0	35.4 ***
Suburban	42.2	42.1	42.4
Small	32.2	36.6	29.1 ***
Medium	42.2	44.1	40.9 ***
Large	25.6	19.3	29.9 ***

a – means and percentages were weighted with the normed panel weight, F1PNLWT
b – these means and percentages were tested between column 2 and 3
*** p < 0.001, ** p < 0.01, * p < 0.05

The dropped sample score slightly lower on test scores, is slightly more

disadvantaged, and contains more minority students compared to the population sample.

Moreover, compared to the population sample, the dropped sample of students attend

more urban and large schools, and have more students attending schools with slightly

higher proportions of students who participate in the free and reduced lunch program. If I

dropped all students and schools with any missing data, the analytic sample would be

Column 2. This sample of students is statistically significantly different than the dropped

sample on all variables, except in the attendance of suburban schools. Based on this

missing data analysis, dropping students with any missing data would bias the results of

this study because the potential analytic sample (Column 2) contains students who are

more advantaged, higher performing, White, and attend smaller, rural, and more

advantaged schools.

Multiple Imputation

In order to deal with the bias in dropping students with missing data, I used

multiple imputation to generate values for the missing data. Multiple imputation is a

method used to simulate missing data based on the existing data and its relationships. For

each set of imputed values, a separate complete dataset is created. The number of

imputations needed depends on the fraction of missing information, but many researchers

in the field suggest that three to 10 imputations are sufficient, the most common being

five (McKnight, McKnight, Sidani, & Figueredo, 2007).

For this study, I used the software IBM SPSS Version 20 to run the multiple

imputation. I imputed the missing information five times, which created five separate

datasets. I then aggregated the five datasets to create one complete dataset to use the

HLM software.[13] By using multiple imputation, I retained an additional 8,265 students

for this analysis.

---

[13] When running analyses with multiply imputed data, software programs that recognize
multiple imputation will produce coefficients for each dataset, and then average the
coefficients together. Since I manually aggregated the five datasets together to create
one, I recognize that I lose some variability in the data. However, in an analysis not
presented here, I ran the same models with nonimputed data and received similar results.

Selected Variables

I used the following variables for my investigation of collective teacher efficacy. This section is divided into student-level variables and school-level variables, and further grouped according to the conceptual model in Chapter 1. I recoded many of these variables from their original form; Appendix B lists the full table of variables. Table 3.2 and 3.3 show the descriptive statistics for these measures.

The majority of students in this sample self-identified as White or Asian students (79.7%), and half of these students were male. Approximately 90 percent of the students in this sample attended public schools, and schools that had some form of tracking (94.6%). The majority of students attended small high schools (65.4%), schools that were located in rural areas (53.6%), and high schools that spanned ninth- through twelfth-grades (55.4%). The average years of teaching experience for this sample of schools is only 5.5 years.

Table 3.2: Student descriptives

Variable	n [a]	M [b]	SD
Dependent Variable			
10th Grade Reading	13739	30.378	9.850
10th Grade Math	13739	43.414	13.700
Prior Achievement			
8th Grade Reading	13739	27.015	8.494
8th Grade Math	13739	36.199	11.699
Socio-economic Status	13739	-0.076	0.744
	n [a]	% [b]	SD
Majority Student	10809	79.7	0.402
Minority Student	2930	20.3	0.402
Male	6828	50.1	0.500

n = 13,739 students
a - n are unweighted
b - means are weighted with a normed panel weight

Therefore, I do not believe that the lack of variability adversely affected the results of this study.

Table 3.3: School descriptives

Variable	n^a	M^b	SD
CTE	820	0.036	0.647
Academic Press	820	-0.002	1.037
Inside Classroom Control	820	0.041	0.564
Outside Classroom Control	820	0.241	0.702
Decision Making	820	0.136	0.735
Principal Leadership	820	0.040	0.708
Sense of Order	820	-0.164	0.693
Collegiality	820	0.131	0.704
School SES	820	-0.114	0.443
Average Years of Teacher Experience	820	5.523 yrs	2.247
Moderating Variable			
Collective Teacher Efficacy*Minority School	820	0.300	0.400
	n^a	$\%^b$	SD
Average % of Students in Remedial Reading	820	8.1	8.692
Average % of Students in Remedial Math	820	7.5	8.530
Tracking in School	795	94.6	0.226
Departmentalization	784	82.5	0.375
Minority School	186	15.1	0.359
School Size			
Small	264	65.4	0.476
Medium	340	26.5	0.441
Large	216	8.2	0.274
Urbanicity			
Urban	248	16.5	0.371
Rural	250	53.6	0.500
Suburban	322	30.0	0.458
Public Schools	717	90.4	0.295
Grade Span of School			
K-12 grade	48	17.2	0.378
6th - 12 grade	96	20.9	0.407
9-12 grade	562	55.4	0.497
10-12 grade	114	6.5	0.247

n = 820 schools

a - n are unweighted

b - means are weighted with a normed school weight

Student-Level Variables

Dependent Variable

Tenth Grade Mathematics and Reading Achievement Scores (*ZF12XRIRR/ZF12XMIRR*):

The Educational Testing Service (ETS) developed a series of cognitive tests covering

mathematics, English, science and history subject areas. The test consisted of 116 items,

21 of which specifically measured reading comprehension. Intended to be completed in

21 minutes, this test measured a students' level of comprehension and interpretation of

five short passages. The mathematics portion consisted of 40 items to be completed in 30

minutes, and examined simple application skills and advanced comprehension and

problem solving questions. ETS administered a different test form for each student,

depending on the base year test scores (Curtin, et al., 2002; Ingels, Scott, Lindmark,

Frankel, Myers, et al., 1992). ETS used Item Response Theory (IRT) to recalibrate

assessment scores in order to compare student scores across different versions of the

assessment (Ingels et al., 1994).

For this study, I focus on only the mathematics and reading tests. In order to

create a more meaningful scale, I standardized both mathematics and reading scores to a

mean of 0 and a standard deviation (SD) of 1, in order for these scores to be comparable

to other standardized scores.

Student-Level Independent Variable of Interest

Race (*MINORITY*): This variable represents the student's race and was recoded into a

dichotomous variable. I combined White and Asian students into one group, and African

American, Latino/a, and Native American/Alaska Native students in another. Ideally,

51

Asian students would be separate from their White counterparts, but because on average, Asian students score similarly to White students on tests such as the SATs and the other minority groups perform similarly to one another, I chose to combine these particular groups together. Other studies, such as Lee & Smith (1996), also group student race groups in this way. I created this variable from the first follow-up race variable (F1RACE) because this variable had less missing data than the base year race variable (BYRACE). All White and Asian students were coded as 0, and African American, Latino/a, and Native American/Alaska Native students were coded as 1. For this sample of students, 20.3 percent are considered to be minorities.

Student Control Variables

Gender (*MALE*): This variable represents the gender of the student participants. All male students were coded as 1, and females were coded as 0. Half of the sample are males (50.1%).

SES (*ZSES*): This variable represents the socioeconomic status (SES) of students based on the original composite variable from the base year questionnaire. I standardized the variable in order for the variable to be comparable to other standardized variables. The mean of this variable is equal to 0 and the SD is equal to 1.

Prior Mathematics and Reading Achievement Scores (*ZBY2XRIRR/ZBY2XMIRR*): This variable represents the prior achievement levels of students before their tenth-grade year. As described above, the NELS:88 student participants took multiple assessments as part

52

of their participation in the survey. These students took tests in four subject areas, but only the mathematics and reading scores were used for this study. As with the tenth-grade scores, the eighth-grade mathematics and reading test scores were standardized to a mean of 0 with a SD of 1.

School-Level Variables

School Structural Variable of Interest

Minority Composition (*MINORSCH*): In order to represent the minority composition within the school, I created a dichotomous variable with the original variable, F1RACE. I recoded students into two groups: one representing "minority" (African American, Latino/a, and Native American/Alaska Native students), and the other representing "non-minority" (White and Asian students). Once regrouped, I aggregated these groups to the school level, and calculated the percentage of minority students within the school. Schools with 40 percent or more African American, Latino/a, and Native American/Alaska Native students were considered to be minority schools, which is consistent with the minority composition variable in Lee and Smith (1996). Of the sample of schools, 15.1 percent are considered to be high minority schools.

School Structural Control Variables

Socioeconomic Status (Z*SCHSES*): This variable represents the socioeconomic status (SES) of the students within the school. I created this variable by aggregating the SES of students within their high school (F1SES), and then standardized it (mean = 0, SD = 1).

Size (*SMALL/MEDIUM/LARGE*): This variable represents the student enrollment based on the composite variable, F1SCENRL. I created three dichotomous variables as follows:

1.   *SMALL*: schools with 0 to 799 students (65.4% of the sample)

2.   *MEDIUM*: schools with 800 to 1,599 students (26.5% of the sample)

3.   *LARGE*: schools with 1,600 or more students (8.2% of the sample)

Sector (*PUBLIC*): This variable represents whether the high school is a public (*PUBLIC* = 1) or non-public school (*PUBLIC* = 0). I created the variable with an existing composite variable (G10CTRL1) that was included in the NELS:88 first follow-up administrator questionnaire. Of the schools in the survey, 90.4% are public schools.

Location (*URBAN/SUBURBAN/RURAL*): This variable represents the urbanicity of the high school. I recoded an existing urbanicity variable (G10URBAN) into three separate variables, each representing their particular location. The variable, *URBAN*, represents schools located in a central city (*URBAN* = 1), *SUBURBAN* represents the area surrounding a central city within a county with a relatively high density population (SUBURBAN = 1), and *RURAL* represents schools beyond the suburban region *(RURAL = 1)*. The majority of schools are in rural areas (53.6%), followed by suburban (30.0%), and urban schools (16.5%).

Grade Span (*ElemHS, MidHS, HS912, HS1012*): This variable represents the grade span of the school. I recoded this variable from F1GSPAN, which was from the administrator

questionnaire. I separated F1GSPAN into four variables: high schools that include

elementary school (*ElemHS*), high schools that include middle school (*MidHS*), high

schools that span ninth- to twelfth-grade (*HS912*), and high schools that span tenth-

through twelfth-grade (*HS1012*). All of these new grad span variables are dichotomous

variables. The majority of schools in this sample span grades nine through 12 (55.4%).

Teacher Experience (*TEACHEXP*): This variable captures the level of teacher experience

in the school. I combined two variables from the NELS:88 dataset: years taught at

elementary school (F1T3_4A) and years taught at secondary school (F1T3_4B). Once I

calculated the total years of teaching experience for each teacher, I averaged the years of

experience by school. Of the schools, the average years of experience is 5.5 years.

School Ability (*RABILITY/MABILITY*): These variables represent the ability level of the

students within the school by measuring the percentage of students who received

remedial reading or mathematics courses. In this sample, 8.1 percent of students were

enrolled in remedial reading courses, and 7.5 percent were enrolled in redial mathematics

courses.

Informal Organization of Schooling Variable of Interest

Factor analysis. All variables used to describe the informal aspects of schooling[14]

were created using exploratory factor analysis. Factor analysis has the ability to identify

related variables and create a more concise representation of an underlying concept (Kim

---

[14] The exception being the moderating variable, which was created by multiplying
collective teacher efficacy and school minority status.

& Mueller, 1978; Pett, Lackey, & Sullivan, 2003). For this analysis, I used principal component analysis with a varimax rotation, which assumes that the information explained by one factor is independent of the information in the other factors (Leech, Barrett & Morgan, 2008). The factors were also rotated so that different items were explained by different underlying factors, and that each factor explained more than one item (Leech et al., 2008).

I used principal component analysis to calculate factors from all 34 variables thought to represent various informal aspects of schooling. This procedure created factor loadings, which is a measure of the contribution an item makes to a particular factor (Blaikie, 2003). Ideally, each variable should have a 'high' loading on only one factor (Blaikie, 2003), which for the purposes of this study was 0.4 or greater. The rotated component matrix is presented in Appendix C. Factor analysis created eight factors using student-level data. In order to create school-level informal schooling variables, I aggregated the composite variables to the school-level to create a variable that was representative of the school and not the individual.

Collective Teacher Efficacy (*CTE*): Of the 12 studies that examine collective teacher efficacy as an organizational property of the school, three use items that ask teachers how they individually feel about certain items ("I" or "you" type questions), whereas 10 studies ask teachers how they feel teachers' in their school feel about certain items ("We" or "teachers" type questions). "I" or "you" type questions include items such as "I sometimes feel it is a waste of time to try to do my best as a teacher" (Newmann, Rutter & Smith, 1989), while "we" or "teachers" type questions include items such as "Teachers

in this school have what it takes to get the children to learn" (Goddard, Hoy & Hoy,

2000). A table of prior literature with the type of efficacy questions is presented in Table

3.4 below.

Table 3.4: Prior literature, collective teacher efficacy scale, and type of items

Study	CTE Scale	# of items	Type of Items
Hoover-Dempsey, Bassler, & Brissie (1987)	Created their own	11	"I" or "you"
Newmann, Rutter & Smith (1989)	Created their own	4	"I" or "you"
Lee, Dedrick, & Smith (1991)	Same as Newmann, Rutter & Smith (1991)	4	"I" or "you"
Goddard, Hoy & Hoy (2000)	Created Collective Efficacy Scale (CES) based on a 21 item-scale from Gibson & Dembo (1984)	21	"we" or "teachers"
Goddard (2001)	CES	21	"we" or "teachers"
Goddard & Goddard (2001)	CES	21	"we" or "teachers"
Hoy, Sweetland & Smith (2002)	12-item version of CES	12	"we" or "teachers"
Goddard, LoGerfo & Hoy (2004)	12-item version of CES	12	"we" or "teachers"
Tschannen-Moran & Barr (2004)	Created their own based on Tschannen-Moran & Hoy (2001) Teacher Sense of Efficacy Scale	12	"we" or "teachers"
Cybulski, Hoy & Sweetland (2005)	12-item version of CES	12	"we" or "teachers"
Adams & Forsyth (2006)	12-item version of CES	12	"we" or "teachers"
Goddard & Skrla (2006)	12-item version of CES	12	"we" or "teachers"

For this study, I use the "I" or "you" type questions to follow the precedent set

from other studies that also examine collective teacher efficacy using general purpose

studies (Newmann, Rutter & Smith (1989) and Lee, Dedrick & Smith (1991)). These items are also identified under the construct of teacher efficacy (see Ingels et al., 1994). These items are:

1.   If I try really hard, I can get through even to the most difficult or unmotivated students (F1T4_5A).

2.   I feel that it's part of my responsibility to keep students from dropping out of school (F1T4_5B).

3.   If some students in my class are not doing well, I feel that I should change my approach to the subject (F1T4_5C).

4.   By trying a different teaching method, I can significantly affect a student's achievement (F1T4_5D).

5.   There is really very little I can do to insure that most of my students achieve at a high level (F1T4_5E).

6.   I am certain I am making a difference in the lives of my students (F1T4_5F) (Ingels et al., 1992b).

I reverse coded item five ("There is really…") in order for all responses to reflect higher scores for higher feelings of teacher efficacy. All six items use a Likert-scale ranging from 1 (strongly disagree) to 6 (strongly agree).

I created a composite variable that represents collective teacher efficacy by employing exploratory factor analysis. Factor analysis produced one factor that included all six collective teacher efficacy variables with a reliability of 0.72.[15] Using principal

---

[15] Cronbach's α measures the reliability of the set of variables, and ranges from 0 to 1. The higher the reliability, the lower the error of measurement (Pett, Lackey, & Sullivan, 2003).

component analysis, 42.56 percent of the total variance in the six items can be explained by the one factor.

Moderating Variable

Collective Teacher Efficacy and Minority Schools (*CTEMINOR*): This variable represents the interaction between collective teacher efficacy and high minority schools. I created this variable by multiplying the two variables together. The mean of this variable is 0.03, with a SD of 0.40.

Informal Organization of Schooling Control Variables

Collegiality (*COLLEGIALITY*): Prior studies have found that staff interaction and a strong sense of collegiality distinguished successful schools from less successful ones (Rosenholtz, 1985; Sweetland & Hoy, 2000; Yasumoto, Uekawa & Bidwell, 2001). Similar to the six-item collective teacher efficacy composite, I created a composite variable that represents teachers' collective feelings of collegiality in their schools. I used factor analysis to create one composite variable from four items that measure this construct. Teachers were asked the following questions on the first follow-up questionnaire.

1. I make a conscious effort to coordinate the content of my courses with teachers in my department/curricular area (F1T4_1A).

2. You can count on most staff members to help out anywhere, anytime – even though it may not be part of their official assignment (F1T4_1B).

3. There is a great deal of cooperative effort among staff members (F1T4_2E).

4.   This school seems like a big family; everyone is so close and cordial (F1T4_2H)

      (Ingels et al., 1992b).

These item responses range from 1 (strongly disagree) to 6 (strongly agree). Factor

analysis extracted one factor that explained 53.72 percent of the total variance in the four

variables. The Cronbach's α was 0.69.

Principal Leadership (*PRINLEAD*): This construct represents teachers' collective

feelings about their principal's leadership in their school. Strong principal leadership is

an essential component of student achievement (Brookover et al., 1979; Cotton, 2003).

Moreover, perceptions of principal's leadership may have a larger effect on Black

students compared to White students (Brookover et al., 1979). Similar to the previous

items, these responses have a six-point, Likert-type scale, with higher scores representing

stronger feelings of agreement. Teachers responded to the following statements:

1.   The principal does a poor job of getting resources for the school (F1T4_1F).

2.   The principal deals effectively with pressures from outside the school that

      interfere with my teaching (F1T4_1C).

3.   The principal sets priorities, makes plans, and sees that they are carried out

      (F1T4_1H).

4.   The principal knows what kind of school he/she wants and has communicated it

      to the staff (F1T4_1O).

5.   The principal lets staff members know what is expected of them (F1T4_2I).

6.   The principal is interested in innovation and new ideas (F1T4_2K).

7.   The principal usually consults with staff members before he/she makes decisions

       that affect us (F1T4_2M) (Ingels et al., 1992b).

I combined these seven variables into one composite variable that represents collective

feelings of leadership within the school. I reverse coded item one ("The principal does a

poor job…") in order for high scores to represent strong, positive feelings about their

principal's leadership. The reliability of these variables were very high ($\alpha = 0.90$), and

this factor explained 62.72 percent of the total variance.

Decision Making Power (*DECISION*): From a questionnaire for 1,500 elementary school

teachers, Moore and Esselman (1994) found feelings of influence and decision making

ability was correlated to reading achievement. For this study, decision making power is a

composite of four items from NELS:88. I drew these items from the teacher

questionnaire, which asked teachers how much actual influence they thought teachers in

their schools had over school policy in each of the following areas:

1.   Determining discipline policy (F1T4_9A).

2.   Determining the content of inservice programs (F1T4_9B).

3.   Setting policy on grouping students in classes by ability (F1T4_9C).

4.   Establishing curriculum (F1T4_9D) (Ingels et al., 1992b).

Answers to this question ranged from "no influence" to "a great deal of influence."

Principal component analysis found this factor to account for 56.07 percent of the total

variance in the four variables. The reliability of these variables was 0.74.

Sense of Control (*CONTROL*): Sense of control and feelings of autonomy are key factors

in student achievement (Cotton, 2003). I created this composite variable based on five

items from the teacher questionnaire. The questionnaire asked teachers to rate 1 through

6 (1 being "no control," and 6 being "complete control") to the following question:

How much control do you feel you have in your classroom over each of the following

areas of your planning and teaching?

1.   Selecting textbooks and other instructional materials (F1T2_17A).

2.   Selecting content, topics and skills to be taught (F1T2_17B).

3.   Selecting teaching techniques (F1T2_17C).

4.   Disciplining students (F1T2_17D).

5.   Determining the amount of homework (F1T2_17E) (Ingels et al., 1992b).

Principal component analysis extracted two composite variables. The first, which

describes "inside classroom control," loaded strongly on items three, four, and five. This

factor explained 46.52 percent of the total variance in the items, and had a reliability of

0.63. The second factor, "outside classroom control," loaded strongly on items one and

two, and explained 21.42 percent of the total variance in the five items. This factor had a

Cronbach's α of 0.72.

Sense of Order (*ORDER*): This variable measures the school's sense of order, including

classroom management. A number of studies found a positive relationship between a

strong sense of order and achievement (e.g., Freiberg, Huzinec, & Templeton, 2009;

Hough & Schmitt, 2011; McGarity & Butts, 2006 I gathered these variables from the

teacher questionnaire administered in 1990. The three variables that were combined through factor analysis were:

1. The level of student misbehavior (e.g., notes, horseplay, or fighting in the halls, cafeteria, or student lounge) in this school interferes with my teaching (F1T4_1E).

2. The amount of student tardiness and class cutting in this school interferes with my teaching (F1T4_1M).

3. Routine duties and paperwork interfere with my job of teaching (F1T4_2A) (Ingels et al., 1992b).

Similar to other items, teachers were asked to respond based on a Likert-scale ranging from 1 to 6, with higher numbers representing stronger feelings of disorder. Factor analysis extracted one factor, which explained 54.17 percent of the total variance of the three variables. The reliability for this variable was 0.61.

Academic Press (*PRESS*): This composite variable measures the academic press of the school. A number of studies cite academic press as an important school characteristic in improving student achievement (e.g., Lee et al., 1995; Lee & Smith, 1999; Lee et al., 1997; Ma, 2003; Phillips, 1997; Sweetland & Hoy, 2000). This composite consists of five variables from the first follow-up school administrator survey and was created utilizing factor analysis. These item responses range from 1 to 5, 1 being "not accurate," 3 being "somewhat accurate," and 5 being "very accurate":

1. Students place a high priority on learning (F1C93B).

2. Classroom activities are highly structured (F1C93C).

3. Teachers at this school press students to achieve academically (F1C93D).

4.   Students are expected to do homework (F1C93E).

5.   Counselors and teachers encourage students to enroll in academic classes

        (F1C93J) (Ingels et al., 1992a).

These five variables had a moderately-high reliability ($\alpha = 0.77$), and explained 51.40

percent of the total variance.

Formal Organization of Schooling Control Variables

Tracking (*TRACKING*): This variable represents the presence of curricular tracking

within the school. The original variable (F1T2_4) asked teachers which of the following

answers best described the achievement level of tenth-graders in the particular class

compared with the average tenth-grader in the school: higher achievement levels, average

achievement levels, lower achievement levels, or widely differing achievement levels. I

used this variable as a proxy for tracking as suggested by Rees, Argys, and Brewer

(1996). I considered schools with classes described as "widely differing achievement

levels" to have no curricular tracking (*TRACKING* = 0), while all other answers confirm

the existence of tracking in the school (*TRACKING* = 1). Almost all schools in the

sample have some form of tracking (94.6%).

Departmentalization (*DEPARTMENT*): Included in the administrator survey, this

variable represents whether the schools' faculty is departmentalized or divided into

subject areas (F1C37). Schools with departments were recoded as 1, while schools

without departmental structures were recoded as 0. The majority of schools in this

sample have departmental structures in place (82.5%).

Weights

Weights are necessary for this type of analysis in order to adjust for the over-sampling of certain groups, such as Asians and Hispanics, for the effects of nonresponse, and to adjust for the variation among schools. For factor analysis and descriptive calculations, I used the first follow-up panel weight, *F1PNLWT*. While the NELS:88 dataset does include a base year school weight, the National Center for Education Statistics advised against using this weight for multilevel analyses (Spencer, Frankel, Ingles, Rasinski & Tourangeau, 1990, as cited in Howley & Howley, 2004). Instead, I use a weight created by Lee and Smith (1995) with the assistance from the Sampling Division of the University of Michigan's Institute for Social Research. They created their own school-level weight based on the "probabilities that students in each school had spent their eighth-grade year in NELS public, Catholic, independent (NAIS), or other private schools, weighted by the total enrollment of each high school" (Lee & Smith, 1995, p. 264). I obtained this weight from Lee and Croninger's HLM methodology course from the University of Maryland, College Park.

Methodology

This study, much like other school-effects research, examines nested data (students within schools) in its exploration of school factors and student-level variables.[16] Using traditional ordinary least squares regression would not account for correlations between students within the same schools, because each school provides relatively similar resources, organizational characteristics, and structural features for all of the

---

[16] This study uses "student-level" and "level-1" interchangeably; likewise for "school-level" and "level-2."

students within their building. Therefore, in order to account for the variation among students and between schools, I employed hierarchical linear modeling, or HLM. HLM not only accounts for the nested nature of these data, but also allows for cross-level interactions. For this study in particular, HLM allows me to examine the effect of collective teacher efficacy (school-level) on minority status (student-level).

Arguably, this study could examine school effects as a three-level model, conceptualized as students within classrooms within schools. I would argue, though, that while most elementary students spend the majority of their day in one classroom, high school students move between multiple classrooms throughout their school day. As such, I believe that the effect of the school, as a whole, would have more of an impact on high school students than their individual classrooms.

Modeling School Variation and Centering Decisions

Many level-1 variables can be transformed, or centered, to allow for more meaningful values. I employ both grand-mean and group-mean centering for this analysis. Grand-mean centering a variable requires the overall mean of the variable to be subtracted from all the values, whereas in group-mean centering, the school mean is subtracted from the value for each student in the school (Hancock, 2010). Grand-mean centering variables accounts for both the individual and school-level effects, whereas group-mean centering only estimates the individual-level effects.

In this study, certain variables, such as student minority status, are tested to see if these measures vary across schools. If these measures do in fact vary, I would be able to run cross-level interactions between student and school variables. If these variables do

not vary or I do not expect these variables to have large differences between schools, I "fix" the slope and assume that the slope is the same for all students.

Moderator Effect

One area of interest is whether collective teacher efficacy has a moderating effect on minority composition of a school and student achievement. Figure 3.1 illustrates this relationship. In order to confirm a moderator, the interaction of the predictor and moderator must be statistically significant (path c). While the main effect of the predictor (path a) and moderator (path b) may or may not be significant, these relationships are not relevant to the test for moderation (Baron & Kenny, 1986).

Figure 3.1: Collective teacher efficacy moderator model

Minority Composition (predictor)

a

Collective Teacher Efficacy (moderator)

b

10th Grade Achievement (outcome variable)

c

Minority Composition
x
Collective Teacher Efficacy (predictor x moderator)
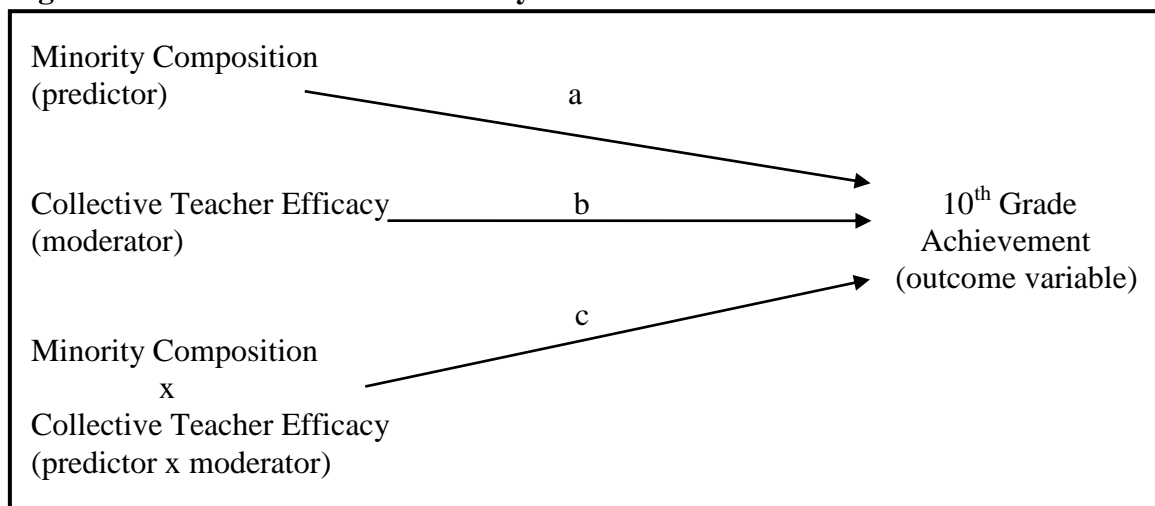
Adapted from Baron & Kenny, 1986

Analytic Strategy

I ran a series of models for this analysis using HLM 7 software. These models were weighted with the school weight described in Lee & Smith (1995, 1996 & 1997). First, I ran a fully unconditional model (FUM) in which no predictors were included in either the level-1 or level-2 equations (see Figure 3.2).

67

Figure 3.2: Fully unconditional model

Level-1
Tenth-grade test score$_{ij}$ = β_{0j} + r_{ij}

Level-2
β_{0j} = γ_{00} + u_{0j}

     where

β_{0j}	=	average math or reading achievement score for all students in school j
r_{ij}	=	error (unmodeled variability) for student i in school j
γ_{00}	=	average math or reading achievement score for all schools
μ_{0j}	=	error (unmodeled variability) for school j

The FUM provides information to calculate the intraclass correlation coefficient (ICC), which is the proportion of between-school variance in the reading and mathematics outcomes (see Figure 3.3). If the ICC is less than 0.05, or less than 5 percent of outcome is due to differences between schools, then there would be little use for using multilevel modeling (Heck, Thomas & Tabata, 2010). As presented in the following chapter, I use HLM for this study since the ICC exceeds the 0.05 benchmark.

Figure 3.3: Intraclass correlation equation

$$ICC = \frac{\tau_{00}}{\tau_{00} + \sigma^2}$$

    where
τ_{00} = between-school variance
$\tau_{00} + \sigma^2$ = total variance

Next, I ran a level-1 model with only student variables in the model. These variables are student gender, race, SES, and eighth-grade test scores. Because I am interested in the variation of the impact of being a minority student on achievement

68

(research question 2), I group-mean centered student minority status, and allowed this measure to vary. I chose to grand-mean center and fix gender, student SES, and eighth-grade prior achievement in order to control for these variables at both levels.

In subsequent models, I added school variables in clusters according to the three school and schooling distinctions and the variables of interest. First, I introduce the three variables of interest into the level-1 model by including collective teacher efficacy, school minority status, and the interaction of both variables. Next, I added variables that represented informal aspects of schooling, followed by measures that represented formal ones. I included school structural variables, such as school sector and SES, last. Because HLM analysis aims for a parsimonious model, I removed variables, aside from the variables of interest, which were not statistically significant at p value ≥ 0.20. Following this decision rule, the final models for the mathematics and reading outcomes are presented in Figure 3.4. These final models were used to answer research questions 1 and 3. The following chapter presents the results from the models described here.

Figure 3.4: Final models

MATHEMATICS:
Level-1
10th GRADE MATH$_{ij}$ = β_{0j} + β_{1j}(MALE$_{ij}$) + β_{2j}(MINORITY$_{ij}$) + β_{3j}(8th GRADE MATH$_{ij}$) + β_{4j}(ZSES$_{ij}$) + r_{ij}

Level-2
β_{0j} = γ_{00} + γ_{01}(MINORSCH$_j$) + γ_{02}(CTE$_j$) + γ_{03}(CTE*MINOR$_j$) + γ_{04}(PRESS$_j$) + γ_{05}(DECISION$_j$) + γ_{06}(COLLEGIALITY$_j$) + γ_{07}(ElemHS$_j$) + γ_{08}(HS912$_j$) + γ_{09}(HS1012$_j$) + γ_{010}(PUBLIC$_j$) + γ_{011}(TEACHEXP$_j$) + γ_{012}(ZSCHSES$_j$) + u_{0j}
β_{1j} = γ_{10}
β_{2j} = γ_{20}
β_{3j} = γ_{30}
β_{4j} = γ_{40}

READING:

Level-1

10th GRADE READING$_{ij}$ = β_{0j} + β_{1j}(MALE$_{ij}$) + β_{2j}(MINORITY$_{ij}$) + β_{3j}(8th GRADE READING$_{ij}$) + β_{4j}(ZSES$_{ij}$) + r$_{ij}$

Level-2

β_{0j} = γ_{00} + γ_{01}(MINORSCH$_j$) + γ_{02}(CTE$_j$) + γ_{03}(CTE*MINOR$_j$) + γ_{04}(PRESS$_j$) + γ_{05}(DECISION$_j$) + γ_{06}(COLLEGIALITY$_j$) + γ_{07}(ElemHS$_j$) + γ_{08}(HS912$_j$) + γ_{09}(HS1012$_j$) + γ_{010}(TEACHEXP$_j$) + γ_{011}(ZSCHSES$_j$) + u$_{0j}$

β_{1j} = γ_{10}

β_{2j} = γ_{20}

β_{3j} = γ_{30}

β_{4j} = γ_{40}

where

β_{0j}	=	average mathematics or reading achievement score for all students in school j
β_{xj}	=	change in achievement for student variable X in school j
r$_{ij}$	=	error (unmodeled variability) for student i in school j
γ_{00}	=	average mathematics or reading achievement score for all schools
γ_{0w}	=	average change in achievement for school variable W
μ_{0j}	=	error (unmodeled variability) for school j

CHAPTER 4: RESULTS

This chapter presents the results from this study's research questions, each dealing with the role of collective teacher efficacy. The first aims to answer the broad effect of collective teacher efficacy on mathematics and reading achievement for tenth-grade students, while the second and third questions deal with the ways it affects minority students and schools – that is, historically disadvantaged students and the schools that they attend.

1.   Does collective teacher efficacy have a positive impact on student achievement?

2.   Does a gap exist between minority students and majority students in regards to achievement within schools? If so, do high schools with high collective teacher efficacy have an equalizing effect on this achievement gap within schools?

3.   Do high schools with large minority concentrations have different achievement levels compared to schools with low minority concentrations? If so, do high schools with high collective teacher efficacy have an interacting or moderating effect on the relationship between schools with high minority populations and student achievement?

I present the results from the various models below. First, I show the results from the fully unconditional model, in order to justify the use of multilevel modeling. Next, I present the results of the model built according to conceptual groups: variables of

interest, informal organization of schooling variables, formal organization of schooling

variables, and school structural variables. If a variable, other than the variables of

interest, are non-significant at $p \geq 0.20$, I remove the variable from the model. Because

the dependent variables are standardized (M = 0, SD =1), coefficients can be interpreted

as effect sizes or a percentage change in the standard deviation (SD) of the dependent

variables associated with a unit change of the independent variable.

Fully Unconditional Models

I examined how much of the variation in the outcomes occurs between schools by

running a fully unconditional model (FUM), which is presented in Table 4.1. Based on

these results, I found that 20.6 percent of the variance in tenth-grade mathematics test

scores, and 12.8 percent of the variance in tenth-grade reading test scores occurred

between schools (Figure 4.1); therefore, it is appropriate to use hierarchical linear

modeling in order to disentangle variance at different levels (i.e., students and schools).

Table 4.1: Fully unconditional models

FUM	Math	Reading
Reliability	0.80	0.69
Average scores across schools, β_0	0.02	0.02
Random Effect		
Intercept, u_0	0.20 ***	0.13 ***
Level-1, r	0.77	0.89

*** p < 0.001, ** p < 0.01, * p < 0.05

Figure 4.1: Intraclass correlation calculations

$$\text{ICC (10}^{\text{th}} \text{ grade mathematics)} = \frac{0.20006}{0.76887 + 0.20006} = 0.206$$

$$\text{ICC (10}^{\text{th}} \text{ grade reading)} = \frac{0.126}{0.85871 + 0.126} = 0.128$$

Tenth-Grade Mathematics Outcome

Student-Level Variables

Table 4.2 presents the level-1 model that only includes the student-level variables on mathematics test scores. I found that, on average, male students scored 0.06 SD lower than females in this sample, and more socially and economically advantaged students scored 0.07 SD higher on this test. Eighth-grade mathematics scores were the largest predictor of tenth-grade mathematics scores – a one SD increase in prior mathematics scores was predicted to increase tenth-grade test scores by 0.76 SD.

In order to answer research question two, I examined whether mathematics test score differences exists between minority and majority students. When controlling for other student characteristics, minority students scored 0.12 SD lower than majority students on the tenth-grade mathematics test, which confirmed the existence of an average achievement gap within schools. In order to examine whether any school variables, including collective teacher efficacy, had an equalizing effect on the achievement gap between minority and majority students, I group-mean centered the variable and allowed this measure to vary across schools. The variance component for the minority slope is not statistically significant ($u_2 = 0.04$, $p > 0.05$); thus, the effect of

being a minority student on mathematics test scores is the same, regardless of the school

that he/she attends. In other words, the difference between mathematics test scores for

minority and majority students does not vary between schools, so it cannot be explained

by school variables. Lastly, the reliability estimates of this model indicate a reliability on

the intercept of 0.57, and very low reliability of the minority slope (0.02).[17] The low

reliability of the minority variable also supports fixing the variable and not allowing it to

vary by schools.

Table 4.2: Mathematics student-level model

Level-1 Mathematics Model		
Reliability Intercept, β_0	0.57	
Minority, β_2	0.02	
Intercept, β_0	0.09	***
Male, β_1	-0.03	
Minority Student, β_2	-0.11	***
Prior Achievement (Grand Mean), β_3	0.83	***
Student SES, β_4	0.07	***
Random Effect	*Variance Component*	
Intercept, u_0	0.02	***
Minority slope, u_2	0.04	
Level-1, r	0.46	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Because prior mathematics test scores explained a large amount of the variation in

tenth-grade test scores, I chose to run the next five models by group-mean centering prior

mathematics test scores. By group-mean centering this measure, I was able to examine

the effects of level-2 variables before controlling for prior achievement.[18] In the final

---

[17] The reliability for this model is similar to the final model because prior achievement
was grand mean centered.
[18] In models not presented here, I ran this analysis with prior achievement fully controlled
(at both the student and school levels) throughout all iterations of model building. When

model, I grand-mean center prior mathematics to determine which, if any, variables have an effect net of students eighth-grade test scores. The results for the remaining mathematics models are presented in Table 4.3 and 4.4. Because the results for the student variables remain relatively consistent, the results I describe below focus on school-level variables.

School Variables of Interest

In Table 4.3 column 1, I introduced the school-level variables of primary interest into the model (the schools minority status, collective teacher efficacy, and the interaction between these variables). This model indicates that collective teacher efficacy had a small, yet positive effect on tenth-grade mathematics scores. I found that a one SD increase in collective teacher efficacy was predicted to increase mathematics test score by 0.10 SD ($p < 0.05$). In addition, high minority enrollment schools performed 0.49 SD lower than low minority enrollment schools on mathematics achievement scores. This model also foreshadows the results for the third research question. I found that the moderating role of collective teacher efficacy, represented by the interaction, was not statistically significant (-0.07 SD, $p > 0.27$).

Informal Organization of Schooling

Table 4.3 column 2 includes all of the informal organization of schooling variables, including feelings of control within and beyond the classroom, academic press, feelings of principal leadership, feelings of decision making abilities, sense of order, and

employing this method, however, I was unable to distinguish whether the school-level variables had any influence on 10th grade achievement because prior achievement explained a large portion of the variance and made almost all school variables non-significant. As such, for the models presented here, I chose to fully control for prior achievement only after I reached a model that was satisfactory for this analysis.

sense of collegiality within the school. I found that increased feelings of academic press and increased feelings of decision making abilities positively impact mathematics test scores, independently (0.13, $p < 0.001$ and 0.09 SD, $p < 0.01$, respectively). Surprisingly, school's collective sense of collegiality had a negative impact, although this effect was very small (-0.06 SD, $p < 0.05$). This may be an indication that collegiality does not always correspond with achievement, especially if the focus of collegiality is not on improving achievement. High minority schools and collective teacher efficacy remained statistically significant.

Formal Organization of Schooling

Table 4.3 column 3 excluded the informal organization of schooling variables that did not meet the $p \geq 0.20$ criteria, and included the formal organization of schooling variables. Neither the existence of tracking nor departmentalization in schools was statistically significant. Academic press, decision making ability, and sense of collegiality remained statistically significant, as did high minority school status and collective teacher efficacy. Because the existence of tracking and departmentalization in schools did not obtain the criteria for inclusion, I remove these variables for the next set of models.

School Structural Variables

School structural variables were added to the model, and presented in Table 4.3 column 4. I found only two variables to have statistical significance: grade span and school SES. High schools that include elementary schools were predicted to score 0.20 SD higher than high schools that spanned grades nine through twelve. Moreover, school

SES was positively and statistically significantly associated with tenth-grade mathematics scores (0.27 SD).

Interestingly, while collective teacher efficacy remained statistically significant at 0.09 SD, the effect of attending a high minority school decreased by more than half when average SES and grade span was included in the model. This result suggests that these structural variables, particularly SES, accounts for some of the variation in high minority schools.[19] The introduction of school structural variables decreased the effect of academic press (0.13 SD to 0.08 SD), and made decision making ability and sense of collegiality non-significant.

---

[19] In a model not presented here, I found that the addition of only school SES to the model presented in column 4 decreased the coefficient of school minority status from -0.43 to -0.25 SD.

Table 4.3: Mathematics HLM models 1-4

Mathematics	1		2		3		4	
Reliability	0.92		0.90		0.90		0.86	
Intercept, β_0	-0.01		-0.03		-0.02		0.04	**
School Variables of Interest								
Minority School, γ_{00}	-0.49	***	-0.41	***	-0.43	***	-0.18	***
Collective Teacher Efficacy (CTE), γ_{01}	0.10	*	0.09	*	0.09	*	0.09	**
CTE*Minority School, γ_{02}	-0.07		-0.08		-0.08		-0.06	
Informal Organization of Schooling Variables								
Inside Classroom Control, γ_{03}			-0.01					
Outside Classroom Control, γ_{04}			0.02					
Academic Press, γ_{05}			0.13	***	0.13	***	0.08	***
Principal Leadership, γ_{06}			-0.01					
Decision Making Ability, γ_{07}			0.09	**	0.09	**	0.05	
Sense of Order, γ_{08}			-0.04		-0.04		-0.02	
Sense of Collegiality, γ_{09}			-0.06	*	-0.06	*	-0.04	
Formal Organization of Schooling Variables								
Tracking, γ_{010}					-0.01			
Departmentalization, γ_{011}					-0.01			
School Structural Variables								
Medium School, γ_{012}							-0.02	
Large School, γ_{013}							0.02	
Urban, γ_{014}							-0.02	
Rural, γ_{015}							0.05	
preK-12+ Grade Span, γ_{016}							0.20	**
6-12+ Grade Span, γ_{017}							0.02	
10-12+ Grade Span, γ_{018}							0.08	
Public School, γ_{019}							0.11	
Ability, γ_{020}							-0.00	
Ave. Years of Experience, γ_{021}							-0.02	
School SES, γ_{022}							0.27	***
Student Variables								
Male, β_1	-0.03		-0.03		-0.03		-0.03	
Minority Student, β_2	-0.11	***	-0.11	***	-0.11	***	-0.12	***
Prior Achievement Group Mean, β_3	0.82	***	0.82	***	0.82	***	0.83	***
Prior Achievement Grand Mean, β_3								
Student SES, β_4	0.09	***	0.09	***	0.09	***	0.07	***
Random Effect	*Variance Component*							
Intercept, u_0	0.16	***	0.13	***	0.13	***	0.09	***
Minority slope, u_2								
Level-1, r	0.21		0.21		0.21		0.21	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Parsimonious Group-Centered Prior Achievement Model

I removed the non-significant variables from the model presented in Table 4.3 column 4 and present these results in Table 4.4 column 5. I found that high minority schools were still predicted to score lower on the mathematics assessment than low minority schools (-0.20 SD, $p < 0.001$). Both collective teacher efficacy and academic press retained their statistical significance, although both coefficients were small. School grade span and school SES also retained their significance. By removing the variables that did not fulfill the criteria, two variables became statistically significant in this model: Decision making ability (0.06 SD) and public school status (0.13 SD).

Final Mathematics Model

I present the results of the final mathematics model in Table 4.4 column 6. This model has a reliability estimate of 0.51, which is substantially lower than the estimate of the preceding models, which ranged from 0.92 to 0.86. Because the final model includes prior achievement as grand-mean rather than group-mean centered, it controls fully for prior mathematics achievement at *both* the student and school levels. Indeed, the lower reliability for the intercept in this model is due to the amount of between school variance in the dependent variable explained by the differences between schools in average eighth grade mathematics achievement (substantially more than any of the prior models in which prior achievement was group mean centered).

The student-level variables remained statistically significant, with the largest effect coming from prior mathematics test scores (0.83 SD, $p < 0.001$). As for school level variables, I found that no variables of interest, no formal organization of schooling variables, and no school structural variables had any impact on tenth-grade mathematics

test scores. This final model answers the main research question (Question 1): contrary to the extant literature, collective teacher efficacy has no statistically significant effect on high school mathematics test scores. In addition, I found that collective teacher efficacy had no moderating role in the relationship between high minority schools and mathematics test score (Question 3).

Although my research questions and hypotheses were not confirmed, a different school variable emerged as an area of interest. Academic press retained its significance across all the models I ran, and although the effect is quite small (0.05 SD, $p < 0.001$), academic press has a greater impact than all other school-level variables, including school SES and minority status. The effect of this school-level variable is smaller than all student-level variables, however. This result suggests that high school student characteristics have a larger effect on mathematics test scores than any school and schooling characteristics, especially students' prior achievement.[20]

---

[20] In a model not presented here, I included an interaction between academic press and high minority schools in the final model. The interaction was not statistically significant.

Table 4.4: Mathematics HLM models 5-6

Mathematics	5		6	
Reliability	0.86		0.51	
Intercept, β_0	0.05	**	0.08	***
School Variables of Interest				
Minority School, γ_{00}	-0.20	***	-0.00	
Collective Teacher Efficacy (CTE), γ_{01}	0.09	**	0.02	
CTE*Minority School, γ_{02}	-0.06		-0.00	
Informal Organization of Schooling Variables				
Inside Classroom Control, γ_{03}				
Outside Classroom Control, γ_{04}				
Academic Press, γ_{05}	0.09	***	0.05	***
Principal Leadership, γ_{06}				
Decision Making Ability, γ_{07}	0.06	*	0.01	
Sense of Order, γ_{08}				
Sense of Collegiality, γ_{09}	-0.03		-0.01	
Formal Organization of Schooling Variables				
Tracking, γ_{010}				
Departmentalization, γ_{011}				
School Structural Variables				
Medium School, γ_{012}				
Large School, γ_{013}				
Urban, γ_{014}				
Rural, γ_{015}				
preK-12+ Grade Span, γ_{016}	0.23	**	0.01	
6-12+ Grade Span, γ_{017}	0.04		0.02	
10-12+ Grade Span, γ_{018}	0.07		-0.00	
Public School, γ_{019}	0.13	*	-0.01	
Ability, γ_{020}				
Ave. Years of Experience, γ_{021}	-0.02		-0.01	
School SES, γ_{022}	0.26	***	-0.01	
Student Variables				
Male, β_1	-0.03		-0.03	
Minority Student, β_2	-0.12	***	-0.11	***
Prior Achievement Group Mean, β_3	0.83	***		
Prior Achievement Grand Mean, β_3			0.83	***
Student SES, β_4	0.07	***	0.07	***
Random Effect	*Variance Component*			
Intercept, u_0	0.09	***	0.01	***
Minority slope, u_2				
Level-1, r	0.21		0.21	

*** p < 0.001, ** p < 0.01, * p < 0.05

Tenth-Grade Reading Outcome

Student-Level Variables

Similar to the initial mathematics model, the level-1 model (Table 4.5) only involves student-level variables. I found that, on average, male students were predicted to score 0.06 SD less than females on the reading test ($p < 0.001$). Students with high SES were also predicted to score better: students with one SD above average SES were predicted to score 0.07 SD higher than those with average SES. Similar to the mathematics model, students who scored well in the eighth-grade reading test were also predicted to score well on the tenth-grade reading test by 0.76 SD. In response to research question two, I found that minority students, compared to majority students, were predicted to score 0.12 SD less within their schools. This result confirms the existence of a gap or difference in tenth-grade reading test scores between historically advantaged and disadvantaged students. This model also demonstrates that this variable does not vary between schools ($u_2 = 0.01$, $p > 0.05$). In other words, all majority and minority students perform similarly, regardless of the school that the student attends. Based on these results, the intercept reliability estimate is moderate (0.55), while the estimate for the minority slope is very low (0.04). This low estimate supports the decision to fix the variance for this measure in the subsequent models.

Once again, because prior achievement explains a large amount of the variation in tenth-grade test scores, I group-mean centered prior achievement for the next five models, in order to examine the school-level effects before controlling for prior achievement. The results for the remaining reading models are presented in Table 4.6

and 4.7. Again, because the results for the level-1 variables remain relatively consistent, the majority of the results I describe below focus on level-2 variables.

Table 4.5: Reading student-level model

Level-1 Reading Model		
Reliability Intercept, β_0	0.55	
Minority, β_2	0.04	
Intercept, β_0	0.07	
Male, β_1	-0.06	***
Minority Student, β_2	-0.12	***
Prior Achievement (Grand Mean), β_3	0.76	***
Student SES, β_4	0.07	***
Random Effect	*Variance Component*	
Intercept, u_0	0.02	***
Minority slope, u_2	0.01	
Level-1, r	0.31	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

School Variables of Interest

Table 4.6 column 1 involves the student-level variables and the primary variables of interest (the schools minority status, collective teacher efficacy, and the interaction between these variables). Of the three school-level variables, I found that only minority school status was statistically significant: compared to schools with low percentages of minority students, schools with high minority enrollments were predicted to score 0.40 SD ($p < 0.001$) lower in tenth-grade reading. Collective teacher efficacy had no statistical effect on tenth-grade reading test scores, nor did it have a moderating role on the relationship between minority schools and achievement. Although both collective teacher efficacy and the interaction were not statistically significant at $p \geq 0.20$, I left these measures in the models because they are the variables of primary interest.

Informal Organization of Schooling

Table 4.6 column 2 includes the informal organization of schooling variables in the model. Similar to the mathematics model 2, academic press had a positive impact on test scores (0.09 SD, $p < 0.001$), as did collective feelings of decision making ability (0.09 SD, $p < 0.01$) and sense of collegiality (-0.06 SD, $p < 0.05$). The effect of attending a high minority school decreased slightly (from -0.40 SD to -0.34 SD), and the other student-level variables remained unchanged from the previous model.

Formal Organization of Schooling

The model in Table 4.6 column 3 removes the informal organization of schooling variables that did not meet the $p \geq 0.20$ criteria, and includes the existence of tracking and departmentalization in the school. Neither of the formal organization of schooling variables was statistically significant; thus, these variables were removed from the model presented in column 4.

School Structural Variables

This model includes school structural variables, such as school size, urbanicity, grade span, school sector, academic ability, average teacher's years of experience, and school SES. Similar to the mathematics model in Table 4.3 column 4, only grade span and school SES were statistically significant. High schools that also include elementary and middle school grades were predicted to score 0.16 SD higher than schools that span grades nine through twelve. Socially and economically advantaged schools (1 SD above the mean) were also predicted to score higher than the average school (0.21 SD, $p < 0.001$).

The addition of school structural variables again lessen the effect of previously significant variables. Minority school status had the largest drop in magnitude – this coefficient decreased in size by half (-0.34 SD to -0.17 SD). Similar to the mathematics model, this result suggests that the effects of school SES and high minority schools are intertwined; a large proportion of the variance associated with the coefficient for high minority schools is the result of minority students being more likely to attend low SES schools.[21]

Parsimonious Group-Centered Prior Achievement Model

This model omits variables that did not meet the criteria for significance and is presented in Table 4.7 column 5. The results from this model were similar to the previous model, although sense of collegiality was no longer statistically significant (-0.04 SD, $p > 0.05$). Although the coefficient for collegiality is the same, the estimate of error increased after excluding the other structural variables from the model.

Final Reading Model

In the fully specified model (Table 4.7 column 6), I fully account for the between-school variation in prior reading achievement by grand-mean centering this measure. I found that no school variables of interest were statistically significant after controlling for differences between schools in average eighth grade reading achievement. This result indicates that collective teacher efficacy has no statistically significant impact on tenth-grade reading achievement (research questions 1 and 3). Furthermore, when fully controlling for prior achievement, high minority schools perform similarly to low

---

[21] In a model not presented here, I found that the addition of only school SES to the model in column 4 decreased the coefficient of school minority status from -0.34 to -0.20 SD.

minority schools. These results also show that collective teacher efficacy has no moderating role in the relationship between high minority schools and reading test scores. The reliability of the final model (0.48) is much lower than the previous five models (ranging from 0.84 to 0.74), which is similar to the mathematics final model. Again, the reduction in reliability between the first model and the final model in Table 4.7 is due largely to controlling for differences between schools in students prior achievement.

As for the other school-level variables, only academic press remained statistically significant. Similar to the mathematics model, student individual characteristics, such as minority status, prior achievement, and SES, had far greater impact than academic press. Although the coefficient is small (0.04 SD, $p < 0.01$), academic press persisted throughout all iterations of the model.[22]

---

[22] In a model not presented here, I included an interaction between academic press and high minority schools in the final model. Similar to the mathematics model, it was statistically non-significant.

Table 4.6: Reading HLM models 1-4

Reading	1		2		3		4	
Reliability	0.84		0.80		0.80		0.74	
Intercept, β_0	-0.00		-0.01		-0.01		0.05	***
School Variables of Interest								
Minority School, γ_{00}	-0.40	***	-0.34	***	-0.34	***	-0.17	***
Collective Teacher Efficacy (CTE), γ_{01}	0.03		0.03		0.03		0.02	
CTE*Minority School, γ_{02}	-0.01		-0.01		-0.02		-0.01	
Informal Organization of Schooling Variables								
Inside Classroom Control, γ_{03}			-0.02					
Outside Classroom Control, γ_{04}			0.00					
Academic Press, γ_{05}			0.09	***	0.09	***	0.06	**
Principal Leadership, γ_{06}			0.00					
Decision Making Ability, γ_{07}			0.09	**	0.08	***	0.05	*
Sense of Order, γ_{08}			-0.04		-0.04		-0.02	
Sense of Collegiality, γ_{09}			-0.06	*	-0.06	*	-0.04	*
Formal Organization of Schooling Variables								
Tracking, γ_{010}					0.02			
Departmentalization, γ_{011}					0.01			
School Structural Variables								
Medium School, γ_{012}							-0.02	
Large School, γ_{013}							0.01	
Urban, γ_{014}							0.01	
Rural, γ_{015}							-0.01	
preK-12+ Grade Span, γ_{016}							0.16	*
6-12+ Grade Span, γ_{017}							0.09	
10-12+ Grade Span, γ_{018}							0.05	
Public School, γ_{019}							-0.01	
Ability, γ_{020}							0.00	
Ave. Years of Experience, γ_{021}							-0.02	
School SES, γ_{022}							0.21	***
Student Variables								
Male, β_1	-0.07		-0.07		-0.07		-0.07	
Minority Student, β_2	-0.11	***	-0.12	***	-0.12	***	-0.13	***
Prior Achievement Group Mean, β_3	0.75	***	0.75	***	0.75	***	0.76	***
Prior Achievement Grand Mean, β_3								
Student SES, β_4	0.09	***	0.09	***	0.09	***	0.06	***
Random Effect	*Variance Component*							
Intercept, u_0	0.11	***	0.09	***	0.09	***	0.06	***
Level-1, r	0.31		0.31		0.31		0.31	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 4.7: Reading HLM models 5-6

Reading	5		6	
Reliability	0.74		0.48	
Intercept, β_0	0.05	***	0.04	***
School Variables of Interest				
Minority School, γ_{00}	-0.16	***	0.00	
Collective Teacher Efficacy (CTE), γ_{01}	0.02		0.00	
CTE*Minority School, γ_{02}	-0.01		0.02	
Informal Organization of Schooling Variables				
Inside Classroom Control, γ_{03}				
Outside Classroom Control, γ_{04}				
Academic Press, γ_{05}	0.06	**	0.04	**
Principal Leadership, γ_{06}				
Decision Making Ability, γ_{07}	0.06	*	0.02	
Sense of Order, γ_{08}				
Sense of Collegiality, γ_{09}	-0.04		-0.01	
Formal Organization of Schooling Variables				
Tracking, γ_{010}				
Departmentalization, γ_{011}				
School Structural Variables				
Medium School, γ_{012}				
Large School, γ_{013}				
Urban, γ_{014}				
Rural, γ_{015}				
preK-12+ Grade Span, γ_{016}	0.17	*	0.09	
6-12+ Grade Span, γ_{017}	0.09		0.04	
10-12+ Grade Span, γ_{018}	0.04		0.01	
Public School, γ_{019}				
Ability, γ_{020}				
Ave. Years of Experience, γ_{021}	-0.02		-0.01	
School SES, γ_{022}	0.21	***	0.01	
Student Variables				
Male, β_1	-0.07		-0.07	
Minority Student, β_2	-0.12	***	-0.11	***
Prior Achievement Group Mean, β_3	0.76	***		
Prior Achievement Grand Mean, β_3			0.76	***
Student SES, β_4	0.06	***	0.06	***
Random Effect	*Variance Component*			
Intercept, u_0	0.06	***	0.02	***
Level-1, r	0.31		0.31	

*** p < 0.001, ** p < 0.01, * p < 0.05

Chapter Summary

This chapter presented results of the research design laid out in the previous chapter, the only difference being that because prior achievement explained a large portion of the variance in test scores, I fully control for prior achievement in the final model only. The three primary research questions for this study are summarized and answered below.

1.   Does collective teacher efficacy have a positive impact on student achievement?

    No, collective teacher efficacy had no effect on either mathematics or reading tenth-grade test scores. Collective teacher efficacy had a small effect on mathematics achievement scores only when prior achievement was not fully accounted for. In the reading models, collective teacher efficacy never gained significance through any iteration of the models.

2.   Does a gap exist between minority students and majority students in regards to achievement within schools? If so, do high schools with high collective teacher efficacy have an equalizing effect on this achievement gap within schools?

    Yes, the data confirmed the existence of an achievement gap in both subjects within schools; however, this variable did not vary between schools. As such, collective teacher efficacy did not have an equalizing effect on this achievement gap.

3.   Do high schools with large minority concentrations have different achievement levels compared to schools with low minority concentrations? If so, do high schools with high collective teacher efficacy have an interacting or moderating effect on the relationship between schools with high minority populations and student achievement?

No, achievement levels in both mathematics and reading did not differ depending on the minority concentration of high schools after controlling for average prior achievement. Thus, collective teacher efficacy had no moderating effect on the relationship between high minority enrollment schools and achievement.

Despite the fact that the variables of interest were not statistically significant, school academic press, an informal organization of schooling control variable, was a statistically significant measure of both mathematics and reading test scores (0.05 SD and 0.04 SD, respectively).

CHAPTER 5: CONCLUSION

The purpose of this chapter is to summarize the previous four sections and to provide a discussion of the study's findings in greater detail. This chapter also examines these results in relation to the current field of literature, reveals limitations in this analysis, and provides suggestions for future research and policy.

Summary and Discussion

This study examined collective teacher efficacy and its impact on student test scores. Based on the school-effects literature, I conceptualized collective teacher efficacy as an informal organizational aspect of schooling. In addition to the informal organization of schools, the formal organization of schools and other school structures were used in this study to provide a more comprehensive model of factors that might influence school effectiveness.

This study answered three research questions on various ways that collective teacher efficacy could impact mathematics and tenth-grade reading test scores:
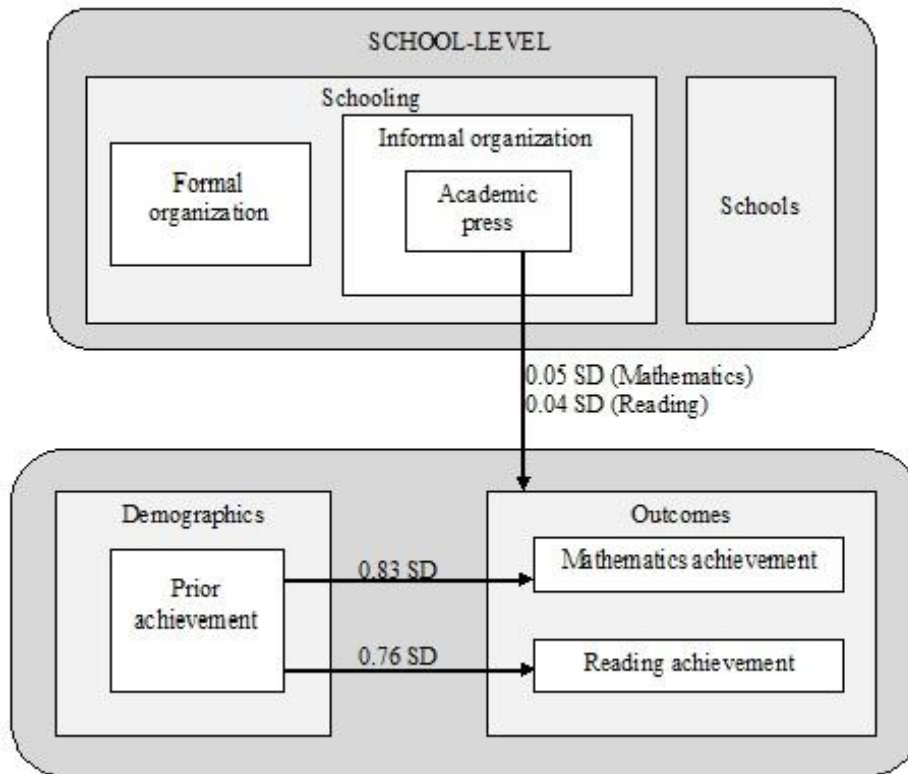
1. Does collective teacher efficacy have a positive impact on student achievement?

2. Does a gap exist between minority students and majority students in regards to achievement within schools? If so, do high schools with high collective teacher efficacy have an equalizing effect on this achievement gap within schools?

3.   Do high schools with large minority concentrations have different achievement

levels compared to schools with low minority concentrations? If so, do high

schools with high collective teacher efficacy have an interacting or moderating

effect on the relationship between schools with high minority populations and

student achievement?

I used the National Education Longitudinal Study of 1988 (NELS:88) to answer

these research questions. These data included a nationally representative sample of

eighth-grade students who were followed into their tenth-grade year. In order to retain

the largest sample size possible, I used multiple imputation to create estimates for

missing data. The final sample for this study involved 13,739 students and 820 schools.

Using hierarchical linear modeling (HLM), I found that 20.6 percent of the variance in

mathematics and 12.8 percent of the variance in reading test scores occurred between

schools, thus supporting the use of HLM for this analysis.

Based on the conceptual model presented in Chapter 1 (Figure 1.2), I built a series

of multilevel models to answer the three research questions. The full models (column 6

on Table 4.4 and 4.7) include all student-level variables and school-level variables that

had statistical significance at $p < 0.20$. The full model helped answer research questions

1 and 3, whereas the student-level model (Table 4.2 and 4.5) addressed research question

2. The results from this study are illustrated in Figure 5.1.

Figure 5.1: Reconfigured conceptual model

Research Question 1: Collective Teacher Efficacy and Test Scores

Mathematics test scores: In the early stages of model building, I found that

collective teacher efficacy had a small, but statistically significant, impact on tenth-grade

test scores (ranged from 0.09 SD to 0.10 SD). However, once I controlled for student

prior achievement scores in the final models, I found that collective teacher efficacy had

no impact on mathematics achievement (0.02 SD, $p > 0.05$). This result conflicted with

the present literature, which found this construct to be an important aspect of schooling.

Reading test score: Unlike the mathematics model, the early stage reading models

did not show collective teacher efficacy as being an important measure of tenth-grade

reading achievement (ranged from 0.02 SD to 0.03 SD). Once I controlled for prior

reading achievement in the final model, I found that collective teacher efficacy, indeed,

had no statistical impact on reading (0.00 SD, $p > 0.05$). Similar to the mathematics model, this result is contrary to the current literature on collective teacher efficacy and reading achievement.

Discussion: Based on the final mathematics and reading models, the results from this study cast doubt about the literature on collective teacher efficacy and achievement. I believe these discrepancies exist due to methodological differences, mainly in population sample, statistical technique, and control variables.

First, of the five articles in the current literature that found a positive relationship between collective teacher efficacy and achievement, two of these studies examined collective teacher efficacy using hierarchical linear modeling: Goddard (2001) and Goddard, Hoy, and Hoy (2000). However, both of these studies examined elementary schools as their sample population and not high schools. Of the two studies that used high schools as their population of interest (Goddard, LoGerfo & Hoy, 2004 and Hoy, Sweetland & Smith, 2002), neither study used HLM as their statistical technique. By using a more sophisticated statistical technique such as HLM, I was able to account for variation at both the student and school levels, which may explain the contradictory results of this study.

Second, I believe this study differs from the extant literature due to the ability to control for prior achievement. Current studies of collective teacher efficacy and achievement in high schools suggest a positive relationship between the two measures, yet neither of these studies controlled for prior achievement. In this study, I found that once prior achievement was accounted for, collective teacher efficacy had no effect on either mathematics or reading test scores. While Goddard (2001) did indeed control for

prior achievement in his HLM model, his population sample was limited to elementary schools in one district located in the Midwest, thus differing in population sample.

As such, I would argue that the results of this study, in conjunction with current research on collective teacher efficacy and achievement, demonstrate that collective teacher efficacy may impact elementary school students only.[23] By the time students reach high school, collective teacher efficacy has no direct effect on student test scores; at best, there is an indirect effect through prior achievement associated with lower grades. The only substantial measure of both the tenth-grade mathematics and reading scores was eighth-grade test scores. This result suggests that by the time students enter high school, high performing students continue to perform well, and students with below average scores will continue to score below average.

Research Question 2: Student Minority Status and Test Scores

The student-level models (Table 4.2 and 4.5) provided insight into possible differences in achievement between majority and minority students, or, stated slightly differently, differences in achievement between historically advantaged and disadvantaged students. For both mathematics and reading tenth-grade test scores, the level-1 model confirmed the existence of an achievement gap (-0.11 SD for mathematics; -0.12 SD for reading). This gap in test scores was not affected by any school-level variables, including collective teacher efficacy.

---

[23] Of the five studies that examined collective teacher efficacy and achievement, two used elementary schools as their sample, two studies examined high schools, and one examined middle schools. While all five studies found statistically significant findings, the elementary school samples were the only studies that used HLM as their quantitative method. I believe that using a more sophisticated technique such as HLM provides more accuracy in results compared to other methods that cannot control for the nested nature of school data.

Discussion: The existence of a test score gap in both mathematics and reading is consistent with other examinations of collective teacher efficacy and achievement. Goddard, Hoy, and Hoy (2000) and Goddard (2001) used minority status as a control and found that African American students scored lower than non-African American students, even after controlling for other student characteristics such as SES and gender. Other studies have also confirmed and examined the existence and persistence of the gap from the time children enter kindergarten, till they enter adulthood (Jencks & Phillips, 1998; Phillips, Crouse & Ralph; 1998).

Though some suggest that by eliminating the differences in young students would reduce the Black-White test score gap at the end of high school (by at least half, if not more), most student and school characteristics do not fully explain this gap (Phillips et al., 1998). Similarly, this study found no school organizational or structural characteristics, including collective teacher efficacy, to affect the achievement gap. Future research should focus on factors that may diminish these differences in order to better create policies and reforms to target this issue.

Research question 3: Minority Schools, Collective Teacher Efficacy and Test Scores

Mathematics test score: In the early stages of the model building, I found that high minority schools scored lower than low minority schools (ranged from -0.49 SD to -0.20 SD). In these early models, the moderating effect of collective teacher efficacy on the relationship between minority schools and achievement (represented as the interaction of minority school and collective teacher efficacy), was not statistically significant. The final model, which controlled for eighth-grade mathematics test scores, showed that

neither high minority schools (-0.00 SD, $p > 0.05$) nor the interaction (-0.00 SD, $p >$ 0.05) had any impact on tenth-grade mathematics achievement.

Reading test score: Similar to the examination of mathematics, the early stage models did not control for prior achievement in order to fully examine the level-2 variables. These early models demonstrated that high minority schools performed statistically significantly lower than low minority schools (ranged from -0.16 SD to -0.40 SD), but the moderating effect of collective teacher efficacy was not statistically significant. The final reading model showed that neither high minority schools (0.00 SD, $p > 0.05$) nor the moderating effect of collective teacher efficacy (0.02 SD, $p > 0.05$) was statistically significant.

Discussion: Both models showed no differences in achievement between schools with high minority enrollments versus low minority enrollments. This result is consistent with other studies, such as Goddard, LoGerfo, and Hoy (2004) and Lee and Smith (1996). When I compare school minority status with earlier models that explain student minority status, this result is notable. Whereas prior achievement could not explain the achievement differences within high schools, it did explain differences between them. This result may be another indication that the academic experiences of students in the lower grades greatly impact the achievement gap between high schools. Not only did eighth-grade prior achievement explain away average minority school status, but it also explained away the effect of school SES. It is likely that the eighth-grade scores are a reflection of earlier effects of schooling from racial and income segregation, especially since many communities have historically had housing policies that excluded low-income families from most communities, including in White and middle class schools (Orfield &

Lee, 2005). The levels of segregation for Black and Latino/a students have also been

increasing since the 1980s (Orfield & Lee, 2005). Because students experience

segregated communities and schools from the time they enter kindergarten, these effects

on achievement seem to accumulate and become so powerful that they wash out other

school effects by the time students are in high school.

Academic Press

One interesting finding of this study was the significance of academic press in both

mathematics and reading test scores (0.05 SD, $p < 0.001$ and 0.04 SD, $p < 0.01$,

respectively). Although the effect was small, it persisted throughout all iterations of the

models, unlike all other school variables in the models. Hoy, Sweetland, and Smith

(2002) also included academic press as a control variable, but unlike this study, found no

significant effect on school mathematics achievement. This discrepancy may be due to

this model controlling for student level variables, whereas Hoy et al. examined school

variables alone.

In this study, academic press was the only level-2 variable that was statistically

significant. It was found to be more important than any other school variables in

impacting mathematics and reading achievement. This result supports other studies that

find academic press to be of import. For example, Phillips (1997), in her examination of

attendance and mathematics, found that various measures of academic press[24] had a

positive effect on school attendance (ranged from 0.09 SD to 0.15 SD), and hours spent

on homework, which Phillips identified as a form of academic press, had a positive

---

[24] Academic press was measured as teachers' expectations of students' high school and
college completion, percent of students in a school that took algebra in the 8th grade, and
the average number of hours per week spent on homework.

relationship with mathematics scores (0.24 SD). Lee et al. (1999) also found academic press to be an important aspect of student learning, but found that it must be in conjunction with strong social supports. These supports included support from teachers, parents, peers, and students' communities. Perhaps the effect of academic press in this study would be stronger if I included these various measures of social support in the models.

I speculate that academic press has an effect unlike other school variables, such as collective teacher efficacy, because this construct taps into activities, policies, or practices that have a direct contribution to achievement. For example, one item that makes up academic press is "students are expected to do homework." Studies have found that students who have been assigned homework performed better on a variety of tests on a variety of subjects, than students who were assigned no homework (Cooper, 2008). The same idea would apply for "counselors and teachers encourage students to enroll in academic classes," because research shows that students who enroll in rigorous classes have more school success (e.g., Gamoran, 1987; Lee, Croninger & Smith, 1997). Collective teacher efficacy, on the other hand, includes items that gauge less tangible behaviors and feelings; items include phrases such as "If I try really hard" and "I feel." While these items are important indicators of teacher beliefs, they do not capture actual activities or expectations that have become institutionalized, such as expecting students to enroll in academic courses.

I also suspect that academic press remained statistically significant because the items that were used to create this construct tapped into feelings about the school, rather than individual feelings. The academic press items were drawn from the school

99

administrator survey, whereas the other informal organization items were drawn from the teacher surveys, which explains why the academic press items are more general and broader in nature. Perhaps the administrator survey provides a more accurate reflection of the informal aspects of schooling, compared to aggregating individualized teacher responses to the school-level to examine the informal organization of schools.

Contributions to the Literature

This study on collective teacher efficacy and achievement contributes to the literature in a number of ways. First, this study examines this relationship using hierarchical linear modeling on a nationally representative sample of eighth-graders entering high school. Past studies that used hierarchical linear modeling involved only elementary schools, while high school studies used different statistical techniques such as a path model. Using my particular sample and statistical technique produced a result that was counter to the previous research on collective teacher efficacy.

Second, the extant literature on collective teacher efficacy had few examinations on the indirect role of this construct on student test scores, and no examinations of collective teacher efficacy as a moderator. Collective teacher efficacy had no moderating effect on high minority schools, because almost all school variables had no statistical effect. Unfortunately, collective teacher efficacy had no effect on reducing the achievement gap between minority and majority high school students. As will be discussed below, continued research is necessary to examine ways in which to reduce the test score gap and the role that policies and practices may play in doing so.

Third, this study contributes to the literature on academic press and achievement. Of all the formal and informal aspects of school organization, academic press was the one and only measure to affect high school test scores. This result is of import because it supports the existing literature on the significance of academic press in schools. Few studies have examined this effect using multilevel modeling, but of those, academic press was found to be an essential aspect of schooling (see Lee et al., 1995; Lee & Smith, 1999; Lee et al., 1997; Ma, 2003). Not only does this result demonstrate the importance of academic press, but it may also indicate the importance of further developing and examining the informal organizational measures in school-effects models. Although effects are small in this study, it was only an aspect of the informal organization of schooling that had a consistent impact on achievement in this study.

Lastly, my conceptualization of school effects research is more comprehensive than previous studies. Although I found fewer significant school effects than other researchers who examine the impact of schools and schooling on achievement, I would argue that this is due to a fuller conceptualization of school-effects research. Not only did I include school and schooling variables (as per Bidwell and Kasarda, 1980), but I further disaggregated schooling by formal and informal organizational measures. By including a more comprehensive picture of schooling into the models, this study may more accurately portray what the effect of these structures in schools around the country.

Limitations and Future Research Directions

This study has a few limitations and corresponding research directions that could be addressed. First, very few school effects were found to be significant in this study of

high schools students. Based on the results from this study, I suspect that larger effects on measures such as collective teacher efficacy may be felt in the earlier grades. Unfortunately, I was unable to examine elementary school students because the youngest students surveyed in the National Education Longitudinal Survey of 1988 were eighth-graders. Future research on collective teacher efficacy could examine younger students, perhaps using ECLS-K: 2011 (which is expected to be available fall 2012), to further investigate the construct using a nationally representative sample of elementary school students.

Future research on collective teacher efficacy should also consider a measure of consensus in the school. Less variability in consensus may more accurately measure a school's collective efficacy, since strong concentrations of teachers with the same feelings of efficacy would reflect strong informal school organization. The research on consensus is limited in numbers and mixed in results, however. Newmann et al. (1989) found a negative relationship between consensus and collective teacher efficacy, while Goddard et al. (2000) found consensus had no statistically significant impact on student achievment. Consensus in collective teacher efficacy may contribute to this literature base on achievement, and may shed light into additional aspects of informal organizing within schools.

Second, while collective teacher efficacy was not significant, this study supports the continued research on academic press. Current studies have found academic press to have an impact on student achievement, and may especially be important for low-income students (Shouse, 1996) and urban elementary school students (Goddard, Sweetland &

Hoy, 2000). The role that academic press, in addition to other informal aspects of schooling, may have on student educational experiences should be further examined.

Third, the data used for this analysis is dated; the first survey year of NELS:88 was conducted over 20 years ago. Many education policies and reforms that affect the classroom have occurred since then, including *No Child Left Behind*. Unfortunately, datasets that included more direct measures of collective and individual teacher efficacy along with achievement scores were not available when this analysis began. I suggest that future surveys include more teacher and school items to better examine this and other informal organization of schooling measures.[25] Likewise, NELS:88 is a general purpose survey and was not created to directly examine collective teacher efficacy. A certain amount of measurement error is associated with using measures to capture constructs that are not purposefully measured by general purpose surveys (Crocker & Algina, 2006).

This last limitation pertains to school effects literature and research, in general. School effects researchers conceptualize and build models differently, and tend to use similar survey items to represent different constructs. For example, Newmann, Rutter, and Smith (1989) and Lee, Dedrick, and Smith (1991) used the same four items from the general purpose survey High School and Beyond – one study used the items to measure teacher efficacy, while the other intended to represent both teacher efficacy and job satisfaction. Moreover, these two studies used teacher responses to "I" or "you" type items to create their collective teacher efficacy measures, whereas other researchers use "we" or "teachers" type items to gauge their school's level of efficacy. Perhaps this inconsistency in the creation of this variable contributes to the discrepancy in results

---

[25] The High School Longitudinal Study of 2009 includes many items that would fulfill this request. Unfortunately, this survey was unavailable at the time of this writing.

between this study and others that examine achievement. Since I chose to follow the precedent set by other general purpose studies (and due to the limited data available in the survey), I aggregated "I" and "you" items to the school level, whereas other studies that examined collective teacher efficacy and achievement used "we" and "teachers" responses to measure collective teacher efficacy. This study demonstrates that the construction of collective teacher efficacy may be more significant than researchers have realized.

Inconsistencies are also not limited to school effects research or to those who use large national datasets. In the initial search for literature on collective teacher efficacy, many studies were eliminated due to differences in concept (i.e., some definitions of collective teacher efficacy were vastly different than the definition used for this study[26]) and measurement (i.e., conceptualized as collective teacher efficacy but measured as individual teacher efficacy[27]). I suggest that future researchers find consistency in the creation and conceptualization of collective teacher efficacy, in order to find consistency in results and interpretations.

Not only is there a lack of consistency in collective teacher efficacy, but school effects researchers represent other constructs using different survey items. For example, school ability could be represented as the average test score from the year prior, or the percentage of students who are proficient on the examination. Many of these differences are due to limitations in available data, but many are decisions left up to the individual researcher. These discrepancies in how we measure certain constructs may change or skew the results and/or the interpretation of our studies.

---

[26] For an example see Ware and Kitsantas (2011)
[27] Some examples include Zambo and Zambo (2008) and Chambers and Cantrell (2008).

Future Policy Directions

Results from this study suggest, at least indirectly, a number of policy directions. First, the strongest predictor of high school test scores was earlier eighth-grade test scores. This may indicate that the achievement trajectories for students are set early and may be difficult to alter by the tenth grade. As such, school reforms that aim to improve student achievement, including reforms to reduce the achievement gap, might be more effective if started early. For example, Success For All (SFA), targets students in grades pre-kindergarten to fifth-grade. Based on a Texas statewide reading evaluation of 111 SFA schools from 1994-1998, the test score gap decreased for both Black and Latino/a students (compared to White students) (Slavin & Madden, 2001). In fact, for those schools that participated the longest (four years), the gap for Black and Latino students was only four and seven percentage points, compared to 14 and 10 percentage points (respectively) for non-SFA schools. While evidence from his study does not fully support a program like SFA, earlier-grade programs may be more effective than programs that target achievement differences in later grades.

Results from this study also suggest the importance of informal aspects of schooling in two policy-related ways. First, current programs and reforms to improve school outcomes should more carefully evaluate ways in which aspects of informal school organization may contribute to the effectiveness of these programs. While this study found a very small effect, it is possible that informal aspects of schools are important prior conditions for the successful implementation of other reforms. Better measures may also identify stronger effects for this aspect of schooling, especially given that there are few agreed upon measures or scales with which to conduct such studies. I

encourage education researchers to better conceptualize the informal aspect of schooling, create scales that appropriately measure this area of research, and to consistently utilize these scales to measure future education interventions and policies.[28]

Second, I suggest that federal education policies more fully take into account the informal aspects of schooling. As we have seen with the increased testing emphasis with *No Child Left Behind*, federal policies can greatly affect what goes on in schools. At the time of this writing, the U.S. Department of Education has yet to formulate any policies that encourage schools and districts to improve student achievement by cultivating positive informal aspects of schooling. A few schools and districts have already recognized this important reform area and have implemented programs such as the Acceptance and Commitment Training project that aims to improve middle school teachers' sense of collegiality and well-being (Oregon Research Institute, 2012). Professional learning communities have also been identified as a method to "engage school staffs in the processes that collectively seek new knowledge and processes" (Mawhinney, Haas & Wood, 2005, p. 11). Greater consideration of these aspects of schools could facilitate the effectiveness of federal and state policies that seek to enhance teacher effectiveness.

Final Thoughts

President Obama recently stated, "A world-class education is the single most important factor in determining not just whether our kids can compete for the best jobs

---

[28] It is also possible that researchers have yet to agree upon a scale because, by definition, informal organizations occur spontaneously and emerge to fulfill a need. As such, measuring the informal organization of a school, in relation to student achievement, may be difficult if the informal organization measures are not constant over time.

but whether America can out-compete countries around the world" (2011). This statement is much like the language in *A Nation at Risk*, written almost 30 years ago, and shows us that many of the same educational concerns exist today as they did in 1983. The federal government has pushed for reforms in accountability, school choice, curriculum standards, and testing to tackle these concerns, yet few policies have aimed to cultivate positive organizational aspects of schooling. Although this study does not provide strong evidence for the effects of the informal aspects of schooling on achievement, I believe that the results from this study warrant further investigation of this area in education as a step towards increasing learning and improving the academic experiences of students in the U.S.

APPENDIX A: MEDIATING EFFECTS

Based on the results from three path analyses, I calculated the mediating, or
indirect, effect and total effect of an independent variable based on Kenny (2011):

$$\text{total effect} = \text{direct effect} + \text{indirect effect}$$
$$c \quad = \quad c' \quad + \quad ab$$
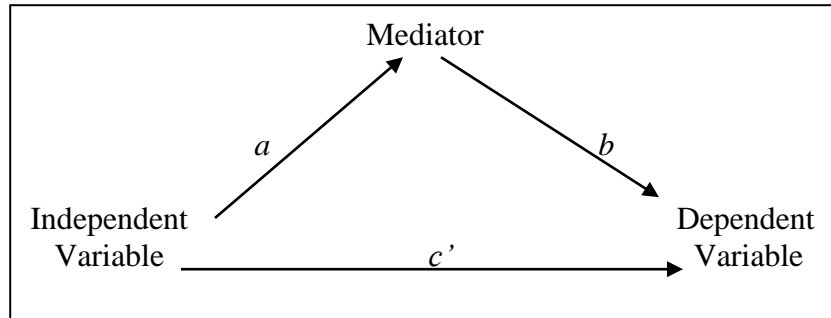
The proportion of the effect that is mediated is calculated as

$$\frac{ab}{c} \text{ or } \left(1 - \frac{c'}{c}\right).$$

According to Kenny, the proportion of the indirect effect should not be calculated unless
the standardized c is at least ± 0.20. In order for a variable to have complete mediation,
this proportion should be at least 0.80.

NELS:88 Variable Name	Description	Construct	Level	New Variable Name
F12XRIRR	10th grade reading IRT-estimated number right	dependent variable	student	ZF12XRIRR
F12XMIRR	10th grade mathematics IRT-estimated number right	dependent variable	student	ZF12XMIRR
BY2XRIRR	8th grade reading IRT-estimated number right	prior achievement	student	ZBY2XRIRR
BY2XMIRR	8th grade mathematics IRT-estimated number right	prior achievement	student	ZBY2XMIRR
RACE	composite race	student race	student	WHITEASIAN
BYSES	SES composite	student SES	student	ZBYSES
SEX	composite sex	student gender	student	MALE
F1T4_5A	I can get through to most difficult student			
F1T4_5B	teachers responsibility to keep students from dropping out			
F1T4_5C	change approach if students not doing well	collective teacher efficacy	school	CTE
F1T4_5D	different methods can affect achievement			
F1T4_5E	little I can do to insure high achievement			
F1T4_5F	teacher making difference in students' lives			
F1C93B	students place high priority on learning			
F1C93C	classroom activities highly structured			
F1C93D	teachers press students to achieve	school's academic press	school	PRESS
F1C93E	students are expected to do homework			
F1C93J	students encouraged to take academic classes			
F1T2_17A	control over text/materials			
F1T2_17B	control over content			
F1T2_17C	control over teaching technique	sense of control	school	CONTROL
F1T2_17D	control over disciplining			
F1T2_17E	control over amount of homework			

F1T4_9A	influence over discipline policy	feelings of influence & decision making	school	DECISION
F1T4_9B	influence over inservice programs			
F1T4_9C	influence grouping students by ability			
F1T4_9D	influence over establishing curriculum			
F1T4_1F	principal poor at getting resources	feelings of principal leadership	school	PRINLEAD
F1T4_1G	principal deals with outside pressures			
F1T4_1H	principal makes plans and carried them out			
F1T4_1O	principal knows what kind of school he wants			
F1T4_2I	principal lets staff know what's expected			
F1T4_2K	principal is interested in innovation			
F1T4_2M	principal consult staff before decisions			
F1T4_1E	student misbehavior interferes w/teaching	sense of order	school	ORDER
F1T4_1M	tardiness/cutting interferes w/teaching			
F1T4_2A	routine duties interfere w/teaching			
F1T4_1A	coordinate course content w/dept teachers	collegiality	school	COLLEGIALITY
F1T4_1B	can count on staff members to help out			
F1T4_2E	great deal of cooperative effort among staff			
F1T4_2H	school seems like a big family			
F1T2_4	achievement level of students in this class compared to other 10th graders in the school	tracking	school	TRACKING
F1C37	is faculty departmentalized?	departmentalization	school	DEPARTMENT
F1SES	10th grade school SES	school SES	school	ZSCHSES
F1RACE	composite race	school racial composition	school	MINORSCH
F1SCENRL	school size	school size	school	SMALL, MEDIUM, LARGE

110

G10URBAN	urbanicity of 10th grade school	school urbanicity	school	URBAN, SUBURBAN, RURAL
F1SGSPAN	grade span of school	grade span	school	ELEMHS, MIDHS, HS912, HS1012
G10CTRL1	school control composite	school sector	school	PUBLIC
F1T3_4A	years taught at elementary level	teacher years of experience	school	TEACHEXP
F1T3_4B	years taught at secondary level		school	
F1C30B	% of students receive remedial reading	school ability	school	RABILITY
F1C30C	% of students receive remedial math	school ability	school	MABILITY
F1T4_5A-F, F1RACE	collective teacher efficacy & high minority school	mediator	school	CTEMINOR
---	school weight as described in Lee & Smith (1995)	weight	school	SCHWEIGHT

APPENDIX C: FACTOR ANALYSIS LOADINGS

NELS:88 Variable Name	Factor Loading	Description	Construct
F1T4_5A	0.664	I can get through to most difficult student	
F1T4_5B	0.585	teachers responsibility to keep students from dropping out	
F1T4_5C	0.69	change approach if students not doing well	collective teacher efficacy
F1T4_5D	0.767	different methods can affect achievement	
F1T4_5E	0.562	little I can do to insure high achievement	
F1T4_5F	0.554	teacher making difference in students' lives	
F1C93B	0.713	students place high priority on learning	
F1C93C	0.631	classroom activities highly structured	
F1C93D	0.828	teachers press students to achieve	school's academic press
F1C93E	0.746	students are expected to do homework	
F1C93J	0.635	students encouraged to take academic classes	
F1T2_17A	0.830	control over text/materials	outside classroom control
F1T2_17B	0.816	control over content	
F1T2_17C	0.741	control over teaching technique	inside classroom control
F1T2_17D	0.679	control over disciplining	
F1T2_17E	0.809	control over amount of homework	
F1T4_9A	0.665	influence over discipline policy	feelings of influence & decision making
F1T4_9B	0.681	influence over inservice programs	
F1T4_9C	0.749	influence grouping students by ability	
F1T4_9D	0.657	influence over establishing curriculum	
F1T4_1F	0.694	principal poor at getting resources	
F1T4_1G	0.725	principal deals with outside pressures	
F1T4_1H	0.839	principal makes plans and carried them out	
F1T4_1O	0.839	principal knows what kind of school he wants	feelings of principal leadership
F1T4_2I	0.837	principal lets staff know what's expected	
F1T4_2K	0.749	principal is interested in innovation	
F1T4_2M	0.640	principal consult staff before decisions	
F1T4_1E	0.737	student misbehavior interferes w/teaching	
F1T4_1M	0.749	tardiness/cutting interferes w/teaching	sense of order
F1T4_2A	0.558	routine duties interfere w/teaching	
F1T4_1A	0.401	coordinate course content w/dept teachers	
F1T4_1B	0.805	can count on staff members to help out	
F1T4_2E	0.812	great deal of cooperative effort among staff	collegiality
F1T4_2H	0.659	school seems like a big family	

REFERENCES

Adams, C. M., & Forsyth, P. B. (2006). Proximate sources of collective teacher efficacy. *Journal of Educational Administration, 44*(6), 625-642.

Armor, D., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., et al. (1976). *Analysis of the school preferred reading program in selected Los Angeles minority schools*. Santa Monica, CA.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117-148.

Bandura, A. (2000). Exercise of human agency through collective efficacy. *Current Directions in Psychological Science, 9*(3), 75-78.

Bankston III, C., & Caldas, S. (1996). Majority African American schools and social injustice: The influence of de facto segregation on academic achievement. *Social Forces, 75*(2), 535-555.

Bankston III, C., & Caldas, S. (1998). Family structure, schoolmates, and racial inequalities in school achievement. *Journal of Marriage and the Family, 60*(3).

Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Bidwell, C. E., & Kasarda, J. D. (1980). Conceptualizing and measuring the effects of school and schooling. *American Journal of Education, 88*(4), 401-430.

Blaikie, N. (2003). *Analyzing quantitative data*. Thousand Oaks, CA: SAGE Publications Ltd.

Borman, G., & Dowling, M. (2010). Schools and inequality: A multilevel analysis of

Coleman's equality of educaional opportunity data. *Teachers College Record, 112*(5), 1201-1246.

Brookover, W., Schweitzer, J., Schneider, J., Beady, C., Flood, P., & Wisenbaker, J. (1978). Elementary school social climate and school achievement. *American Educational Research Journal, 15*(2), 301-318.

Cantrell, S. C., & Hughes, H. K. (2008). Teacher efficacy and content literacy implementation: An exploration of the effects of extended professional development with coaching. *Journal of Literacy Research, 40*(1), 95-127.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, New Jersey: Lawrence Erlbaum Associates

Coleman, J. S., United States, O. o. E., & Statistics, N. C. f. E. (1966). *Equality of educational opportunity*. Washington: U.S. Dept. of Health, Education, and Welfare, Office of Education.

Cooper, Harris. (2008). *Homework: What the ressearch says*. Reston, VA: The National Council of Teachers of Mathematics.

Cotton, K. (2003). *Principals and student achievement: What the research says*. Alexandria, VA: Association for Supervision and Curriculum Development

Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Independence, KY: Wadsworth

Curtin, T. R., Ingels, S., Wu, S., Heuer, R., RTI, & Owings, J. (2002). *Quick guide to using the NELS:88/2000 data*. Washington, DC.

Cybulski, T. G., Hoy, W. K., & Sweetland, S. R. (2005). The roles of collective efficacy

    of teachers and fiscal efficiency in student achievement. *Journal of Educational*

    Administration, 43(5), 439-461.

Dearing, E., Kreider, H., Simpkins, S., & Weiss, H. (2006). Family involvement in

    school and low-income children's literacy performance: Longitudinal associations

    between and within families. *Journal of Educational Psychology*, 98, 653-664.

Diamond, J. B., Randolph, A., & Spillane, J. (2004). Teachers' expectations and sense of

    responsibility for student learning: The importance of race, class and

    organizational habitus. *Anthropology & Education Quarterly, 35*(1), 75-98.

Entwisle, D., Alexander, K., & Olson, L. (2005). First grade and educational attainment

    by age 22: A new story. *American Journal of Sociology, 110*(5), 1458-1502.

Evans, A. (2009). No Child Left Behind and the quest for educational equity: The role of

    teachers' collective sense of efficacy. [Article]. *Leadership & Policy in Schools,*

    8(1), 64-91.

Ferguson, R. (1998). Can schools narrow the Black-White test score gap? In C. Jencks &

    M. Phillips (Eds.), *The Black-White test score gap*. Washington, DC: The

    Brookings Institution.

Freiberg, H., Huzinec, C. & Templeton, S. (2009). Classroom management − A pathway

    to student acheievment: A study of fourteen inner-city elementary schools.

    Elementary School Journal, 110(1), 63-80.

Gamoran, A. (1987). The stratification of high school learning opportunities. *Sociology of*

    Education, 60(3), 135-155.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*(4), 569-582.

Goddard, R. D. (2001). Collective efficacy: A neglected construct in the study of schools and student achievement. *Journal of Educational Psychology, 93*(3), 467-476.

Goddard, R. (2002). A theoretical and empirical analysis of the measurement of collective efficacy: The development of a short form. *Educational and Psychological Measurement, 62*(1), 97-110.

Goddard, R. D., & Goddard, Y. L. (2001). A multilevel analysis of the relationship between teacher and collective efficacy in urban schools. *Teaching and Teacher Education, 17*(7), 807-818.

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

Goddard, R., Hoy, W. K., & Hoy, A. W. (2004). Collective efficacy beliefs: Theoretical developments, empirical evidence, and future directions. *Educational Researcher, 33*(3), 3-13.

Goddard, R. D., LoGerfo, L., & Hoy, W. K. (2004). High school accountability: The role of perceived collective efficacy. *Educational Policy, 18*(3), 403-425.

Goddard, R. D., & Skrla, L. (2006). The influence of school social composition on teacher's collective efficacy beliefs. *Educational Administration Quarterly, 42*(2), 216-235.

Goddard, R. D., Sweetland, S. R., & Hoy, W. K. (2000). Academic emphasis of urban

elementary schools and student achievement in reading and mathematics: A

multilevel analysis. *Educational Administration Quarterly*, 36(5), 683-702.

Goldschmidt, P., & Wang, J. (1999). When can schools affect dropout behavior? A

longitudinal multilevel analysis. *American Educational Research Journal, 36*(4),

715-738.

Hancock, G. R., & Mueller, R. O. (Eds.). (2010). *The reviewer's guide to quantitative*

methods in the social sciences. New York: Routledge

Hanson, E. M. (2003). *Educational administration and organizational behavior* (5th ed.).

Boston: Pearson Education, Inc.

Hanushek, E. (1997). Assessing the effects of school resources on student performance:

An update. *Educational Evaluation and Policy Analysis*, 19(2), 141-164.

Hanushek, E., & Rivkin, S. (2009). Harming the best: How schools affect the black-white

achievement gap. *Journal of Policy Analysis and Management, 28*(3), 366-393.

Heck, R., Thomas, S., & Tabata, L. (2010). *Multilevel and longitudinal modeling with*

IBS SPSS. New York: Taylor and Francis Group, LLC.

Heckman, J., & Krueger, A. (2005). *Inequality in America*. Cambridge: The MIT Press.

Hoover-Dempsey, K., Bassler, O., & Brissie, J. (1987). Parent involvement:

Contributions of teacher efficacy, school socioeconomic status, and other school

characteristics. *American Educational Research Journal, 24*(3), 417-435.

Hough, D. & Schmitt, V. (2011). An ex post facto examination of relationships among

the developmental designs professional development model/classroom

management approach, school leadership, climate, student achievement,

attendance, and behavior in high poverty middle grades schools. *Middle Grades Research Journal*, 6(3), 163-175.

Howley, C., & Howley, A. (2004). School size and the influence of socioeconomic status on student achievement: Confronting the threat of size bias in national data sets. *Education Policy Analysis Archives*, 12(52). Retrieved from http://epaa.asu.edu/epaa/v12n52/.

Hoy, W. K., & Miskel, C. (2008). *Educational administration: Theory, research, and practice* (8th ed.). New York: McGraw-Hill.

Hoy, W. K., Sweetland, S. R., & Smith, P. (2002). Toward an organizational model of achievement in high schools: The significance of collective efficacy. *Educational Administration Quarterly, 38*(1), 77-93.

Ingels, S., Dowd, K., Baldridge, J., Stipe, J., Bartot, V., Frankel, M., et al. (1994). *National Education Longitudinal Study of 1988: Second follow-up: Student component data file user's manual*. Washington, DC.

Ingels, S., Scott, L., Lindmark, J., Frankel, M., Meyers, S., NORC, et al. (1992a). *National Education Longitudinal Study of 1998: First follow-up: School component data file user's manual*. Washington, DC.

Ingels, S., Scott, L., Lindmark, J., Frankel, M., Myers, S., NORC, et al. (1992b). *National Education Longitudinal Study of 1988: First follow-up: Teacher component data file user's manual*. Washington, DC.

Ingels, S., Thalji, L., Pulliam, P., Bartot, V., Frankel, M., NORC, et al. (1994). *National Education Longitudinal Student of 1988: Second follow-up: Teacher component data file user's manual*. Washington, DC.

Jencks, C., & Phillips, M. (1998). The Black-White test score gap: An introduction. In C.
Jencks & M. Phillips (Eds.), *The Black-White test score gap*. Washington, DC:
The Brookings Institution.

Jussim, L., Eccles, J. S., & Madon, S. (1996). Social perception, social stereotypes, and
teacher expectations: Accuracy and the quest for the poweful self-fulfilling
prophecy. *Advances in Experimental Social Psychology, 28*, 281-388.

Kenny, D. (2011). *Mediation*. Retrieved October 28, 2011, from
http://davidakenny.net/cm/mediate.htm#IE

Kim, J. & Mueller, C. (1978). *Introduction to factor analysis: What it is and how to do it*.
Newbury Park, CA: SAGE University Paper.

Ladd, H., & Ludwig, J. (Eds.). (2003). *The effects of MTO on educational opportunities
In Baltimore: Early evidence*. Washington, DC: Urban Institute.

Lee, S. (1996). *Unraveling the "Model Minority" stereotype*. New York: Teachers
College Press

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case
of school effects. *Educational Psychologist, 33*(2), 125-141.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high
school achievement. *Sociology of Education, 62*(3), 172-192.

Lee, V. E., Croninger, R., & Smith, J. B., (1997). Course-taking, equity, and mathematics
learning: Testing the constrained curriculum hypothesis in U.S. secondary
schools. *Education Evaluation and Policy Analysis,* 19(2), 99-121.

Lee, V. E., Dedrick, R. F., & Smith, J. B. (1991). The effect of the social organization of
schools on teachers' efficacy and satisfaction Sociology of Education, 64, 190-208.

Lee, V. E., & Smith, J. B. (1995). Effects of high school restructuring and size on early

  gains in achievement and engagement. *Sociology of Education*, 68(4), 241-270.

Lee, V. E., & Smith, J. B. (1996). Collective responsibility for learning and its effects on

  gains in achievement for early secondary school students. *American Journal of*

  Education, 104(2), 103-147.

Lee, V. E., & Smith, J. B. (1997). High school size: Which works best and for whom?

  Educational Evaluation and Policy Analysis, 19(3), 205-227.

Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents

  in Chicago: The role of school academic press. *American Educational Research*

  Journal, 36(4), 907-945.

Lee, V. E., Smith, J. B., & Croninger, R. (1995). Another look at high school

  restructuring: More evidence that it improves student achievement, and more

  insight into why. *Issues in Restructuring Schools*, (9), 1-10.

Leech, N., Barrett, K. & Morgan, G. (2008). *SPSS for Intermediate Statistics, 3rd Edition*.

  New York: Lawrence Erlbaum Associates

Lew, J. (2006). "Burden of acting neither White nor Black: Asian American identities in

  context. *The Urban Review*, 38(5), 335-352.

Liu, X. (2007). The effect of teacher influence at school on first-year teacher attrition: A

  multilevel analysis of the Schools and Staffing Survey for 1999-2000. Educational

  Research and Evaluation, 13(1).

Ma, X. (2003). Sense of belonging to school: Can schools make a difference? *The*

  Journal of Educational Research, 96(6), 340-349.

Mawhinney, H., Haas, J., & Wood, C. (2005). *Teachers' perceptions of collective*

efficacy and school conditions for professional learning. Paper presented at the
Annual Meeting of the University Council for Educational Administration.

McGarity, J. & Butts, D. (1984). The relationship among teacher classroom management behavior, student engagement, and student achievement of middle and high school science students of varying aptitude. *Journal of Research in Science Teaching*, 21, 55-61.

McMurrer, D., & Sawhill, I. (1998). *Getting ahead: Economic and social mobility in America*. Washington, DC: The Urban Institute.

McKnight, P., McKnight, K., Sidani, S., & Figueredo, J. (2007). *Missing data: A gentle introduction*. New York: Guilford Press

Mellnik, Ted. (2012). Census figures: Minority babies now in majority. *USA Today*.

Moore, W. P., & Esselman, M. E. (1992). *Teacher efficacy, empowerment, and a focused instructional climate: Does student achievement benefit?* Paper presented at the Annual Conference of the American Educational Research Association. Retrieved from
http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED350252&site=ehost-live.

Moore, W. P., & Esselman, M. E. (1994). *Exploring the Context of Teacher Efficacy: The Role of Achievement and Climate*. Paper presented at the Annual Conference of the American Educational Research Association.

Moretti, E. (2007). Crime and the costs of criminal justice. In C. Belfield & H. Levin (Eds.), *The Price We Pay* (pp. 142-159). Washington, DC: The Brookings Institution.

Muenning, P. (2007). Consequences in health status and costs. In C. Belfield & H. Levin (Eds.), *The Price We Pay* (pp. 125-141). Washington, DC: The Brookings Institution.

Newmann, F. M., Rutter, R. A., & Smith, M. S. (1989). Organizational Factors that Affect School Sense of Efficacy, Community, and Expectations. *Sociology of Education, 62*(4), 221-238.

Orfield, G. & Lee, C. (2005). *Why segregation matters: Poverty and education inequality*. Cambridge, MA: Harvard University

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

Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.

Pajares, F. (2002). Overview of social cognitive theory and of self-efficacy. Retrieved from http://des.emory.edu/mfp/eff.html.

Pett, M., Lackey, N., & Sullivan, J. (2003). *Making senes of factor analysis*. Thousand Oaks: Sage Publications, Inc.

Phillips, M. (1997). What makes schools effective? A comparison of the relationships of communitarian climate and academic climate to mathematics achievement and attendance during middle school. *American Educational Research Journal*, 34(4), 633-662.

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the Black-White test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The Black-White test*

score gap. Washington, DC: The Brookings Institution.

Obama, Barack. (2011). Retrieved on February 27, 2012 from

http://www.whitehouse.gov/issues/education

Oregon Research Institute. (2012). *Teacher and staff wellbeing*. Retrieved on February

27, 2012 from http://teacherstaffwellbeing.com/

Rees, D., Argys, L., & Brewer, D. (1996). Tracking in the United States: Descriptive

statistics from NELS. *Economics of Education Review, 15*(1), 83-89.

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*.

Washington, DC: Economic Policy Institute.

Rouse, C., & Barrow, L. (2006). U.S. elementary and secondary schools: Equalizing

opportunity or replicating the status quo? *The Future of Children, 18*(2), 99.

Rumberger, R., & Palardy, G. (2005). Does segregation still matter? The impact of

student composition on academic achievement in high school. *Teachers College

Record, 107*(9), 1999-2045.

Santos, J. R. (1999). Cronbach's alpha: A tool for assessing the reliability of scales.

Journal of Extension, 37(2).

Shouse, R. (1996). Academic press and sense of community: Conflict, congruence, and

implications for student achievement. *Social Psychology of Education*. 1(1), 47-

68.

Slavin, R. & Madden, N. (2001). *Reducing the gap: Success for All and the achievement

of African-American and Latino students*. Paper presented at the Annual

Conference of the American Educational Research Association. Retrieved April

20, 2012 from

http://www.successforall.org/SuccessForAll/media/PDFs/Reducing_the_AA_and_
Latino_Gap.pdf

Somers, C., & Piliawsky, M. (2004). Drop-out prevention among urban, African
American adolescents: Program evaluation and practical implications. *Preventing
School   Failure, 48*(3), 17-22.

Spencer, B., Frankel, M. Ingels, S., Rasinski, K., & Tourangeau, R. (1990). *National
Educational Longitudinal Study of 1988: Base year sample design report*.
Washington, DC: National Center for Education Statistics.

Sweetland, S. R., & Hoy, W. K. (2000). School characteristics and educational outcomes:
Toward an organizational model of student achievement in middle schools.
Educational Administration Quarterly, 36, 703.

Tschannen-Moran, M., & Barr, M. (2004). Fostering student learning: The relationship of
collective teacher efficacy and student achievement. *Leadership and Policy in
Schools, 3*(3), 189-209.

Tucker, C. M., Porter, T., Reinke, W. M., Herman, K. C., Ivery, P. D., Mack, C. E., et al.
(2005). Promoting teacher efficacy for working with culturally diverse students.
Preventing School Failure, 50(1), 29.

U.S. Department of Education. (2011). *Digest of Education Statistics, 2010*. Washington,
DC.

U.S. Department of Education. (2010). *ESEA Reauthorization: A blueprint for reform*.
Washington DC.

U.S. Department of Education. (1983). *A nation at risk*. Retrieved from
http://www.ed.gov/pubs/NatAtRisk/recomm.html.

U.S. Department of Education. (2009). *Race to the Top Executive Summary*. Washington, DC.

U.S. Department of Education. (n.d.). National Education Longitudinal Study of 1988. Retrieved on October 1, 2011 from http://nces.ed.gov/surveys/nels88/.

U.S. Department of Education, National Center for Education Statistics. (2011). *Digest of Education Statistics, 2010*. Washington, DC.

Ware, H., & Kitsantas, A. (2007). Teacher and collective efficacy beliefs as a predictors of professional commitment. *Journal of Educational Research, 100*(5), 303-310.

Ware, H., & Kitsantas, A. (2011). Predicting teacher commitment using principal and teacher efficacy variables: An HLM approach. *The Journal of Educational Research*, 104, 183-193.

Weber, M. (1909-1920). The bureaucratic machine. In C. Lemert (Ed.), *Social Theory* (3rd ed.). Boulder: Westview Press.

Woolfolk, A. E., & Hoy, W. K. (1990). Prospective Teachers' Sense of Efficacy and Beliefs about Control. *Journal of Educational Psychology*, 82(1), 81-91.

Zambo, R., & Zambo, D. (2008). The impact of professional development in mathematics on teachers' individual and collective efficacy: The stigma of underperforming. *Teacher Education Quarterly, 35*(1), 159-168.