The Effect of Working Conditions on Teacher Effectiveness: Value-added Scores and Student Perception of Teaching

Yincheng Ye

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Educational Research and Evaluation

Kusum Singh, Chair
Bonnie Billingsley
M. David Alexander
Gary Skaggs

May 4th, 2016
Blacksburg, VA

Keywords: School Working Conditions, Effective Teaching, Teacher Value-added Score, High-Need Schools

The Effect of Working Conditions on Teacher Effectiveness: Value-added Scores and Student Perception of Teaching

Yincheng Ye

ABSTRACT

This dissertation presents a quantitative study of the effects of multiple aspects of working conditions on teacher effectiveness as measured by value-added scores and student perceptions of teaching. The data were derived from the 2009-2010 Teacher Working Condition Survey and Student Perception Survey in Measures of Effective Teaching (MET) Project. Using the structural equation modeling and other related methods, several models of teacher effectiveness were estimated. The results supported that instruction and classroom related working conditions at school played important role in effective teaching and student achievement gains in English language arts and mathematics. It was found that, after controlling for teachers' education degree and experience, instructional practice support had significant effect on teachers' value-added scores. Moreover, Classroom autonomy and support for student conduct management were found to have indirect effect on teacher value-added score mediated through the students' perceptions of teaching. In addition, student perceptions of teaching was found to be significantly worse in high-need schools than schools serving fewer minority students or students from low-incoming families, but teacher value-added score was not significantly different between the high versus low needs schools. The findings of the study significantly contributed to a better understanding of the effects of working

environment and how these are related to teacher performance. The study has both theoretical and practical significance; it provided critical evidence that can be used by policy makers to promote teachers' performance, especially in high-needs schools.

Acknowledgements

First of all, I would like to gratefully and sincerely thank my advisor, Dr. Kusum Singh, for her generous time and commitment. I appreciate the support and encouragement she provided for my academic study. Dr. Singh guided me how to do research and helped improve my academic writing, and shared a lot of research experience and insight with me during my graduate studies.

I also would like to express my sincere appreciation to Dr. Bonnie Billingsley, Dr. M. David Alexander, and Dr. Gary Skaggs, for serving as my committee members and giving me invaluable academic advice whenever needed.

Finally, I would like to thank my family for the love, support and constant encouragement that they gave to me over the years. In particular, I would like to thank my husband and my father, for their support as well as some academic suggestions. Their love and support has helped me throughout my doctoral journey, and will help me make progress in the future.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

Background of the Study

The research has clearly shown that effective teachers are key to student learning. Over the past several decades, having effective teachers has been consistently identified as the most important school-based factor in improving students' academic achievement (Aaronson, Barrow, & Sander, 2007; Lockwood & McCaffrey, 2009; Rivkin, Hanushek, & Kain, 2005b; Rockoff, 2004). Moreover, researchers have carefully tracked students' achievement over time and identified that teacher effects have long-run consequences for their students' success (Rivkin, Hanushek, & Kain, 2005a; Rowan, Correnti, & Miller, 2002). The role of the teacher involves more than simply standing in front of a classroom and lecturing, it aims to assist students with making connections and therefore better learning through an educational process in an integrated teaching and learning environment. In other words, an effective teacher understands that teaching involves multiple tasks to ensure that all students receive a quality education (Markley, 2006).

Prior studies have substantiated that teachers differ in their impact on student learning and achievement (Brophy & Good, 1986; Guarino, Hamilton, Lockwood, & Rathbun, 2006; Nye, Konstantopoulos, & Hedges, 2004; Rowan, Correnti, & Miller, 2002; Sanders & Horn, 1995; Wayne & Youngs, 2003; Xue & Meisels, 2004). Teachers are identified as an important source of variability in student achievement (McCaffrey, Lockwood, Koretz, & Hamilton, 2003; Rockoff, 2004).

Recent reform efforts have two main goals that are aimed at closing student achievement gaps and ensuring high-quality education. For both of these goals, effective teachers are the most critical link. Thus, policy makers and practitioners are interested in exploring ways to improve teacher effectiveness. As spurred by Race to the Top, policy makers have undertaken a wide range of reforms and developments to improve the performance of teachers. While in recent years, researchers have consistently shown that effective teachers are distributed very unevenly among schools, especially to the clear disadvantage of high-need schools (Clotfelter, Ladd, & Vigdor, 2011; Lankford, Loeb, & Wyckoff, 2002; Sass, et al. 2012). Many attempts have been made to establish an equitable distribution of effective teachers among schools through national or state educational policies (Race to the Top Progress, 2013).

Although recruiting knowledgeable and skilled teachers is important, it is insufficient for schools to ensure effective teaching performance (Berry, Daughtrey, & Wieder, 2009; 2010). Good teachers need a workplace that promotes their efforts in a variety of ways to retain their effective teaching and doing their best work with students. Teacher effectiveness is not just about teachers' experience, knowledge, skills, etc.; but also about the conditions under which they work. Jackson (2014) concluded that "Teachers may be more or less effective as a contextual function of schools' working conditions that transpose human capital into productivity and effective instructional practice of teachers". (p. 8) Teachers' working conditions play an important role in a school's ability to deliver high quality education. Schools that are able to offer their teachers a safe, pleasant, and supportive working environment can better attract and retain good teachers and even motivate them to do their best. Generally, it covers a broad range of factors and issues, from working time, security to remuneration, as well as the physical conditions and mental demands that exist in schools.

Teacher effectiveness

Teacher effectiveness is conceptualized as the joint function of what it contributes to student achievement outcome and what the teachers do in classrooms (Goe, Bell, & Little, 2008). Despite common perceptions, effective teachers cannot reliably be identified only based on teacher credentials, certification status, or teaching experiences. The best way to assess teachers' effectiveness is to look at their on-the-job performance, including what they do in the classroom and how much progress their students make on achievement tests. In addition to helping in increasing the student achievement score, effective teachers should have the ability to incorporate various methods of teaching and instruction, for a better delivery of the knowledge in classroom such as the use of technology, online resources, etc. (Markley, 2006; Vogt, 1984).

The research work of Sanders and Wenglinsky demonstrated that teacher effectiveness can be measured based on student test scores and are critical to student success (Sanders & Horn, 1995; W. Sanders & J. Rivers, 1996; Sanders, Wright, & Horn, 1997; Wenglinsky, 2000). Teacher value-added model (VAM) is one of the prominent methods to measure teachers' impact on their student achievement, which captures the pure student achievement gains by controlling for other factors that affect achievement, such as individual ability, family environment, past schooling, etc (Aaronson et al., 2007; Glazerman et al., 2010; Hanushek, 1971; Kane & Staiger, 2008; Nye, Konstantopoulos, & Hedges, 2004; Rivkin et al., 2005; Rockoff, 2004).

In addition, previous studies also demonstrated that student perception of teaching can be used as a complement to measures of teacher effectiveness in student learning (Kyriakides, 2005; Oesterle, 2008; Peterson, Wahlquist, & Bone, 2000; Wilkerson, Manatt, Rogers, & R., 2000). Effective teaching meant teachers help their students to understand the content knowledge and engage in learning through the classroom instructions and interactions with their students. Teaching is a reciprocal connection between teachers and students, and students are central to the work of teachers (Follman, 1992, 1995). Students' perceptions of teaching can provide more information and a more robust definition of teacher effectiveness, and are worth considering for inclusion in teacher evaluation systems based on consistent findings of the previous research (Bill & Melinda Gates Foundation, 2010a; 2012; Goe et al., 2008).

Teachers' working conditions

Schools provide a working environment and professional community for teachers by making sure appropriate teaching assignment; enough access to information, materials and technology; and adequate time to work with colleagues on matters of instruction (Little, 1993). There are alternative ways to conceptualize different aspects of teacher working conditions (Ingersoll, 2001; Johnson 2006; Leithwood, 2006; Perie & Baker, 1997). Many factors contribute to working conditions that can make teachers be more effective and help their students to achieve (Johnson, 2003). Researchers have examined the impact of school level working conditions, such as school facilities, community relations (Johnson, Kraft, & Papay, 2012; Ladd, 2011; Loeb, Darling-Hammond, & Luczak, 2005); as well as classroom-instruction level working conditions, such as the amount of instruction support offered and time allotted for instruction planning and collaboration (Johnson et al., 2012; Ladd, 2011). Research has shown the importance of instruction related working conditions for teachers' growth and students' success, because these factors are more directly linked to the classroom instruction and teaching (Johnson, 1990; 2006; Johnson et al., 2012; Leithwood, 2006). These multiple aspects of working conditions are malleable and dynamic within a rich, professional context that encourages teachers' learning and growth. When the schools provide a series of supports for classroom instruction as a good working environment, teachers are more sustained and effective in their work (Johnson et al., 2012; Loeb et al., 2005).

Problem Statement

In the past, discussions of teacher working conditions have focused primarily on teacher salaries and benefits, class size, and internal transfer policies, as well as other similar issues addressed in traditional contractual arrangements (Ladd, 2011; Loeb, Darling-Hammond, & Luczak, 2005; Sass, Hannaway, Xu, Figlio, & Feng, 2012). In addition to these typical issues, it is also meaningful to understand important working conditions, the ones that matter to meaningful teacher effectiveness of both teaching in classroom and student outcomes. However, there is little agreement about which working conditions matter most. Sykes (2008) concluded that "Absent specification provided by theory or a model, causal relations — among working conditions and teacher and student outcomes — cannot but remain murky and unresolved." (p.2)

Despite growing recognition of the importance of working conditions to teacher effectiveness, researchers have only begun to understand the importance of the classroom and instruction level working conditions on teachers' effectiveness. In most studies, working conditions are either operationalized as a single variable or multiple aspects of working conditions are analyzed (Jackson, 2014; Johnson et al., 2012; Ladd, 2011); few studies have focused on the overall and detailed aspect of classroom and instruction level working conditions. Several earlier researchers have focused on investigation of the effect of classroom-instruction level working conditions such as instructional professional development, teaching workload, etc. on teacher effectiveness (Cohen & Hill, 2001; Jackson, 2014; Rosenholtz, 1989; Wei, Darling-Hammond, Andree, Richardson, & Orphanos, 2009); but these studies did not examine the overall effect of classroom and instruction level working conditions and compare the difference in each aspects' effects simultaneously. Therefore, there is a need to get a deeper understanding of the classroom-instruction level working conditions, and fill the gap in understanding how they differ in affecting teacher effectiveness.

Moreover, the research focused on multiple measures of teacher effectiveness that incorporated student perception about teaching and teacher value-added scores is limited. Many studies were restricted to whole-school achievement measures and rarely used controls for past test performance and student background. Instead of teacher value-added measures or student perceptions in teaching, the dependent variables have often been achievement levels. The use of students' perceptions about classroom teaching as a measure of teacher effectiveness is also limited. Most literature focused on the teacher effectiveness in classroom teaching has based on a small sample of studies, by using qualitative methodologies, and thus, has limited generalizability. In a synthesis study on how working conditions that matter to teacher effectiveness, Berry (2010) concluded that more finely tuned research still needs to be conducted to gauge the most critical working conditions linked to effective teaching and student achievement gains. Besides, most empirical work has investigated the impact of teacher perceived working conditions on teacher effectiveness by using state administrated data (Johnson et al., 2012; Ladd, 2011; Loeb, Darling-Hammond, & Luczak, 2005); and few has incorporated the national wide data sets across different states.

In addition, since more recent research has consistently shown that inequality in distribution of effective teachers among schools is to the clear disadvantage of high-need schools (Clotfelter, Ladd, & Vigdor, 2011; Lankford, Loeb, & Wyckoff, 2002); and previous research has found empirical evidence that teachers' effectiveness is much stronger in low-need schools than in high-need schools (Boyed, et al., 2008; Sass, et al., 2012). However, in exploring the teacher effectiveness gap, most studies only investigated the difference in teacher characteristics and background qualities, but did not investigate how the working conditions contribute to the teacher effectiveness gap. The research focused on disparities in the overall working conditions and specific aspects of working conditions that affect teacher effectiveness among schools is limited. Not all aspects of working conditions were considered by previous researchers, such as the instructional supports, or student conduct management.

Therefore, it is necessary to conduct more comprehensive research which takes into consideration both teacher value-added scores and student perceptions of teaching as measures of teacher effectiveness, and how the overall classroom-instruction level working condition, as well as each aspect of working conditions, contribute to the variability in teacher effectiveness, particularly in high-need compared to more advantaged schools. Advancing our understanding of these relationships is particularly important, as this can give rich information about which working condition factors and school characteristics warrant special attention in order to help attract and develop effective teachers.

Purpose of the Study

This study aims to assess the relationships of a broad range of teachers' working conditions and measures of teacher effectiveness. To better understand the impact of working conditions on teacher effectiveness, a conceptual model of working condition and teacher effectiveness was developed based on previous research and theory. More specifically, the conceptual model included how teachers' working conditions affect teacher value-added scores and students' perceptions of teaching, controlling for teacher background qualities. In addition to the overall instruction and classroom level working conditions, the effects of each working condition was examined separately. The study tested and compared the extent to which each working condition contributes to the variation in teacher value-added score and student perceptions of teaching. It was also hypothesized that students' evaluation of teaching could be a mediator between certain aspects of working conditions and teacher effectiveness as measured by value added scores. Furthermore, the study examined whether all constructs and relationships vary across high-need versus low-need school context. Specifically, the study aims to address the following specific research questions:

1) What are the relationships among teacher value-added score, student perceptions about teaching, and teacher background quality in elementary and middle public schools?

2) What is the effect of each of the working conditions (i.e. instructional practice support, teaching workload, instructional resources, classroom autonomy, and support for managing student conduct) on teacher value-added score and student perceptions of teaching, and which work conditions have stronger effects on teaching in elementary and middle public schools?
3) How do the impacts of working conditions on teacher value-added score and student perceptions of teaching differ according to high-need school context (i.e. minority, free and reduced lunch)?

The research questions were addressed with the analysis of secondary datasets, from 2010-2011 Measures of Effective Teaching (MET) study.

Conceptual Model

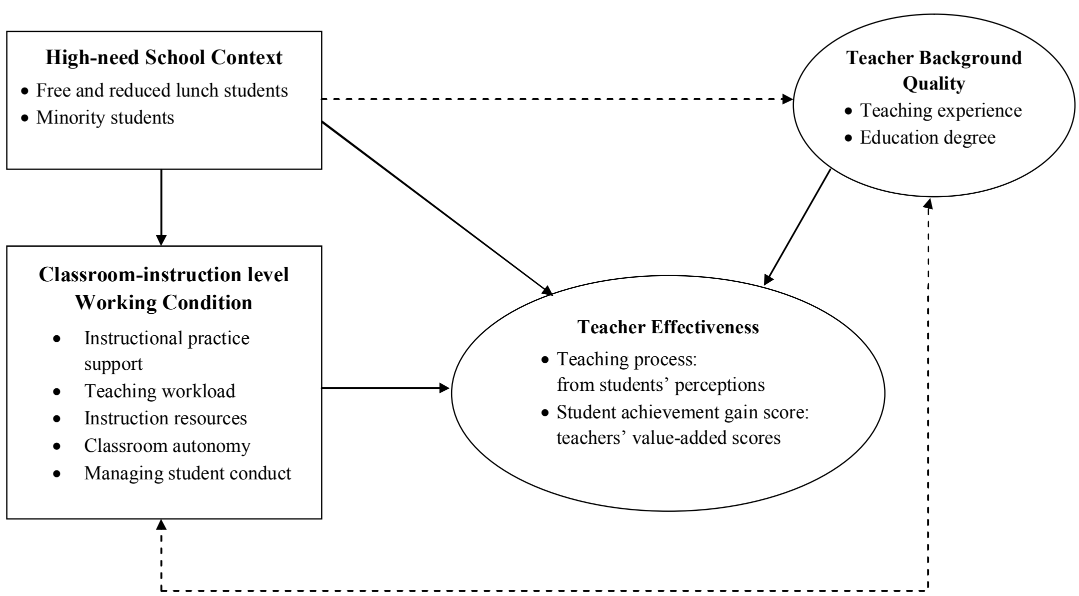

Figure 1.1 Conceptual Model

The conceptual model (displayed in Figure 1.1) proposes that teacher effectiveness is influenced by working conditions, school context, and teacher background qualities. On the right side of the model, the conceptual model allows for the possibility that teacher effectiveness is influenced by teachers' background qualities such as teaching experience, and advanced education degree. Although observed or measurable skills and knowledge appear weakly related to effectiveness, prior research does suggest a link (Clotfelter, Ladd, & Vigdor, 2007; Rockoff, et al. 2008). The arrow linking the teacher background quality and teacher effectiveness addresses the first research question. The boxes on the left side of the model depict the school-level variables of interest. The conceptual model hypothesizes that teachers' working conditions influence teachers' ability to teach effectively and thus influence students' achievement scores. The working conditions include five classroom-instruction level working conditions: instructional practice support, teaching workload, instructional resources, classroom autonomy, and support for managing student conduct. The arrow linking the working conditions and teacher effectiveness addresses the second research question. In addition, the conceptual model also represents the hypothesis that contextual factors in high-need school as defined by student composition may influence teacher effectiveness that addresses the last research question. It is also hypothesized that working conditions and school contextual factors are related through a set of relationships (the last research question). Besides, the dashed arrows represent the possibilities that there are correlations between teacher background quality and school related factors. For example, schools try to attract and recruit good teachers, and teachers are more likely to retain in or move to better schools.

Research Methodology

The research was designed to identify the relative importance of working conditions (i.e. managing student conduct, instructional support, teaching workload, instruction resource, and classroom autonomy) on teacher value-added score and student perceptions of teaching, and to examine whether school context plays a role in differences in these constructs. To achieve the above goals, this research proposed to use a quantitative approach that integrates structural equation modeling (SEM) and related statistical analysis. SEM is a statistical method that takes a confirmatory approach to data analysis of a structural theory bearing on the relationships of some variables of interest (Byrne, 1998). SEM is an especially appropriate method for analyzing non-experimental data. Moreover, SEM deals with the relationships between latent variables, which are free of random error (Loehlin, 1987).

Several statistical steps were taken to prepare the data for the main SEM analyses. First preliminary and descriptive analysis was used for examining the data and assessment of the pattern of correlations among variables. Second the analytic strategy used exploratory factor analysis (EFA) on a subset of the data and confirmatory factor analysis (CFA) to identify the internal structure of the main constructs---teacher working conditions and student perception based on implicit theoretical framework. Then all subsequent relationships were estimated among teacher working conditions, student perceptions, teacher value-added scores, and teacher background quality using SEMs. Additionally, a cross-validation technique was used for validating the estimates of SEMs. Finally, the multi-group analysis incorporating invariance testing of SEMs was conducted in order to test whether interrelationships among constructs differ across subgroups specified by high-need school context. The design of the research analysis can bring empirical evidence to support the conceptual model.

Significance of the Study

This study has both theoretical and practical significance; it strengthens the current theory on the role of work conditions on teacher performance. It is likely to uncover the relationship between teaching effectiveness and teachers' working conditions that directly related to instruction and classroom management from multiple perspectives. First, the proposed study incorporates overall and each aspect of working conditions in the model to explore the differences in their impact on teacher effectiveness. Moreover, the study includes the measures of teacher effectiveness from two domains: student achievement outcomes and students' satisfaction about teaching practices. Often the teacher effectiveness models were either focused student achievement outcomes (teacher value-added scores), or they focused on student evaluations of their teachers' teaching process (student perception). The model in current study presents a more complex view of teacher effectiveness by hypothesizing the relationships between the two domains of teacher effectiveness and working conditions.

In addition, the research is expected to provide critical evidence that can be used by policy makers or researchers to promote teachers' performance. Much of what has been uncovered about the conditions under which teachers teach and how to improve them is important for policymakers and practitioners to consider. The variations of teacher effectiveness among schools have been a major issue across the country. It leads to the inequality of educational opportunity and unequal resources for students' learning.

In order to address teacher effectiveness, many states and districts have implemented policies to enhance physical working conditions, professional development opportunities, and instructional programs. In developing such programs, policy makers would benefit from findings of this study that is likely to provide the relative importance of working condition factors affecting teacher value-added and their student perception toward teaching. The study highlights the differential impact of the various conditions on learning outcomes.

Furthermore, this study integrates theories and background knowledge from several different disciplines, including workplace learning and motivation to enrich the exploration of an important problem in education. This interdisciplinary approach brings new light and an alternative framework for the investigation of how different working conditions contribute to teacher effectiveness.

In terms of methodology, this study provides more accurate estimates of the relationships among factors by using a structural equation modeling approach rather than the regression type analyses. The SEM approach considers both measurement error and interrelationships among factors in a comprehensive analysis. The cross-validation method examines the model robustness and confirms the accuracy of model estimates.

Organization of the Study

In chapter 2, a comprehensive literature review related to teacher effectiveness and working conditions is presented. Specifically, the theories and empirical studies on the relationships between teacher effectiveness and working conditions, and teacher background quality, in English language arts and mathematics are presented. A summary of the reviewed literature gap is presented at the end of chapter 2. In chapter 3, the detailed research methods are described as following: the information of the data source, sample, measures of the variables, primary used analysis method (SEM), and each analysis procedure. In chapter 4, the results followed by the analysis procedures are presented. First, research methods are described as following the descriptive and correlation results of each variable are presented. Second, the measurement testing results of exploratory factor analysis and confirmatory factor analysis are described. Next, the structural model results corresponded to each research question are presented. Then the results of cross-validation of measurement models and structural models are presented follow. Lastly, the invariance testing of structural model results between high-need and low-need schools are presented. In chapter 5, the findings are summarized and the implications are discussed. Finally, the limitations of this research and recommendations for future research are addressed.

CHAPTER 2

LITERATURE REVIEW

This study explored the relationships between teacher working conditions and two measures of teacher effectiveness: student perceptions of teaching and teacher value-added scores. Literature relevant to the research was presented in this chapter in the following sections: (1) definition of teacher effectiveness; (2) two measures of teacher effectiveness: student perceptions of teaching and value-added scores; (3) definition of working conditions; (4) working conditions and teacher effectiveness; (5) each aspect of working conditions and teacher effectiveness: managing student conduct, instructional practice, teaching workload, instruction resource, and classroom autonomy; and (6) differences in these factors between high-need and low-need schools.

Definition of Teacher Effectiveness

Research offers plenty of definitions of an effective teacher. Based on a comprehensive review of the literature, Goe et al. (2008) grouped conceptualizations of teacher effectiveness into three distinct steps: inputs, processes, and outputs (Goe, Bell, & Little, 2008). Teacher inputs included what a teacher brings to their position. Examples of inputs could be teacher knowledge, certification or amount of experience. Researchers have found the teacher inputs influence student learning (Clotfelter, Ladd, & Vigdor, 2007; Darling-Hammond, 2006; Darling-Hammond & Sykes, 2003; Hanushek, 1971; Rivkin, Hanushek, & Kain, 2005). For example, in a scientifically-based research study, Darling-Hammond & Youngs (2002) pointed out that teacher inputs, such as teacher preparation and certification, matter for student achievement. Studies using data involving teacher, school, district, or state levels reported significant relationships between teacher qualifications and student performance in reading or mathematics. Moreover, in Clotfelter, Ladd and Vigdor's (2007) study, they used a rich administrative data set from North Carolina to explore the relationship between teacher characteristics, credentials and student achievement. Their findings indicated that a teacher's experience, regular licensure, and credential, which measures teacher preparation and knowledge of teaching and learning, all have positive effects on student achievement. The effects were found to have a stronger correlation for math than for reading. However, investigators also indicated that most of these measures on teacher inputs, such as teacher advanced educational degree, teacher certification, and teaching experience, have a relatively weak relationship to students' learning outcomes (Aaronson, Barrow, & Sander, 2007; Rivkin et al., 2005). In an influential study that examined the importance of teachers in Chicago public high schools using matched student-teacher administrative data, Aaronson, Barrow, and Sander (2007) showed that after controlling for student and classroom backgrounds, none of the teachers' qualities, including certification, advanced degrees, quality of college attended, and undergraduate major had statistically significant effects on student achievement. Similarly, Rivkin, et al. (2005) also indicated that little of the variation in student achievement gain was explained by observable teacher characteristics such as education degree or teaching experience. Additionally, in a synthesis study, Goe (2007) concluded that teacher licensing for teaching and an education degree are positively correlated with student achievement, but only in mathematics. However, for social studies, science and other important school subjects, significant associations with student achievement have not been found.

Teacher process refers to what happens in the classroom between the teacher and student. Most research of teacher process focused on teaching instruction and classroom activities, such as a teacher planning rigorous lessons that engage students; and the relationships between students and teachers in a related aspect of teaching in the classroom (Goddard, Tschannen-Moran, & Hoy, 2001; Sook-Jeong, 2007). Vogt (1984) stated that effective teachers should have the ability to provide differential instruction to students according to their different abilities. They should further consider different learning modes for their students and match the instruction to the student needs. In Collins's (1990) work with the Teacher Assessment Project, five criteria were established for an effective teacher: (1) is committed to students and learning, (2) knows the subject well, (3) is responsible for managing students, (4) can think systematically about their own practice, and (5) is involved in the learning community. For example, good teachers motivate students' desire to learn and encourage them to be responsible for their own learning. Good teachers should also have the ability to: make instruction accessible to all students, promote students' higher-order thinking skills through effective classroom discussions, questioning, and learning tasks, and clarify with students' learning intentions/targets for success. To be worth noted, teaching is a reciprocal way between teachers and students, and students are central to the work of teachers. Effective teaching process means that teachers help their students to understand the content knowledge and engage them in learning through the classroom instruction and interactions with their students. Research has pointed out the importance of interaction between teachers and their students in supporting instruction in classroom; an effective teacher and student interactions were able to “teach students to think, provide ongoing feedback and support, and facilitate language and vocabulary” (Muntner, 2008; Pianta, et al., 2008).

Research on teacher outputs pertains to the impact of teaching on student outcomes such as student academic achievement. Obviously, the definition of effective teachers refers to someone who can increase student knowledge and enable students’ learning (Clark, 1993). While identifying effectiveness of teachers in student achievement is complicated by the fact that teachers often have very different students, there are several ways to analyze student outcomes. One type of statistical model utilizes a regression based procedure which is common when investigating the impact of student or school level variables on student achievement (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; R. Goddard et al., 2001; Goldhaber & Brewer, 1997; Sook-Jeong, 2007). The research work of Sanders and Wenglinsky has demonstrated that teacher effectiveness can be quantifiable based on student test scores and was critical to student success (Sanders & Horn, 1995; W. Sanders & J. Rivers, 1996; Sanders, Wright, & Horn, 1997; Wenglinsky, 2000). Using student-teacher matched achievement data, these studies have found that student achievement gains were much more influenced by a student's assigned teacher than other factors like class size and class composition. Similarly, Chetty, Friedman, and Rockoff (2011) confirmed Sanders’s conclusions and showed the significant long-term impact of teacher effectiveness measured by student achievement gains on student success. For example, students taught by more effective teachers had better college entrance, higher salaries, higher living community SES, etc. after their graduation (Chetty, Friedman, & Rockoff, 2011).

Multiple Measures of Teacher Effectiveness

As the nation's attention is increasingly focused on the assessment of teaching by using multiple methods, teacher effectiveness should not be defined by single measures. It is meaningful to base it on multiple measures. Papanastasiou (1999) stated "that no single teacher attribute or characteristic is adequate to define an effective teacher" (Papanastasiou, 1999). Multiple methods are suggested to be used to identify effective teachers (Bill & Melinda Gates Foundation, 2010b). In a synthesis study, Markley (2006) suggested that the best ways to measure teacher effectiveness are using an approach that combines teacher classroom observation with student-teacher matched achievement data. In his opinion, an effective teacher is "one who demonstrates knowledge of the curriculum, provides instruction in a variety of approaches to varied students, and measurably increases student achievement." (p. 9)

Teacher input, process, and output are all important measures of teacher effectiveness, but recently, the focus has moved from teacher inputs to teacher process and outputs in terms of students' achievement and success (Darling-Hammond & Youngs, 2002; Goe, 2007; Rice, 2003; Wilson & Floden, 2003). For example, in a synthesis study, Goe (2007) concluded that defining teacher effectiveness solely through teacher input qualifications is not sufficient for ascertaining teacher quality. The background qualifications, such as those certifications, cannot always predict which teachers will be most successful in the classroom. Teacher process in teaching and outputs are better measurements of what is most important: "what the best teachers know and do that results in greater student learning in the classroom." (p. 46). After NCLB, the accountability pressures intensified, the emphasis became on more objective measures such as student outcomes. However, the research focused on multiple measures of teacher effectiveness that incorporated student perceptions of teaching and student achievement gains is limited.

In addition, research also asserted a strong link between student perceptions about teaching and value-added scores as measures of teacher effectiveness (e.g. Hanover Research, 2013; Raudenbush, 2013). Wilkerson, et al. (2000) conducted a study of nearly 2,000 K-12 students and found that student ratings of teaching in both reading and mathematics were significantly more accurate in predicting student achievement than teacher's self-ratings, principal ratings, and principal summative ratings. Most recently, the Bill & Melinda Gates Foundation's Measures of Effective Teaching (MET) Project piloted the Tripod Survey with more than 3,000 school teachers in six large school districts throughout the United States. In their findings, students perceive ratings in teaching process were clearly different among teachers and were predictive of student achievement gain scores. Teachers with more satisfied student survey results had significantly better value-added scores. This study indicated that "students seem to know effective teaching when they experience it." (Bill & Melinda Gates Foundation, 2012a, p. 2)

Teacher Value-added Score

Teacher value-added models are a prominent method to measure a teacher's impact on his/her student achievement, which captures the pure student achievement gains by controlling for other factors that affect achievement, such as individual ability, family environment, past schooling, etc (Aaronson et al., 2007; Glazerman et al., 2010; Hanushek, 1971; Kane & Staiger, 2008; Nye, Konstantopoulos, & Hedges, 2004; Rivkin et al., 2005; Rockoff, 2004). Value-added models are promising, controversial, and increasingly common as a method for determining teacher effectiveness (McCaffrey et al, 2003). The research of Sanders (Sanders, 1996, 1999; Sanders, Wright, & Horn, 1997) and others at the University of Tennessee demonstrated that value-added scores can be a valid measure of teacher effectiveness. Their work asserted that teacher effectiveness is the single biggest contributor to student success. Teacher effectiveness outweighs all other factors, such as class size, classroom context, socioeconomic status, etc. Sanders and Rivers (1996) also conducted a longitudinal study by using the student achievement data from the Tennessee Value-Added Assessment System database to run multivariate analyses of students who took the Tennessee Comprehensive Assessment Program test. The results showed that the teacher effectiveness quantified by value-added score appropriately captured teachers' effects on student achievement. The differences in student achievement of 50 percentile points were observed as a result of teacher sequence after three years. Sanders, Wright and Horn (1997), who followed up the original work of Sanders and Rivers (1996), examined the relative magnitude of teacher effects on student achievement while simultaneously considering the influences of intraclassroom heterogeneity, and class size on student academic growth. The results indicated that teacher effects were dominant factors affecting student academic growth and that the classroom context variables of heterogeneity among students and class sizes have a relatively weak influence on academic growth. Wenglinsky (2000), building on the work previously conducted by Sanders and others, used the data from the eighth grade science report of the National Assessment of Educational Progress (NAEP) to estimate the teacher value-added scores. He found that classroom practices, such as the use of small-group instruction or hands-on learning, and professional development, are all influential factors to student achievement gains. The most significant of the areas was classroom practices. Darling-Hammond (2000) studied data from the Schools and Staffing Surveys and the NAEP data to identify teachers' effects on student achievement score in standard tests. The results indicated that students who are assigned to ineffective teachers have significantly lower achievement and smaller gains in mathematics and reading on NAEP assessments. This study also indicated that states, such as North Carolina, that invested heavily in improvements to teacher effectiveness showed the greatest achievement gains in continues years.

Furthermore, researchers also examined the reliability and validity of using value-added scores as measures of teachers' causal impacts on student achievement. Kane and Rothstein (2008) examined the teacher value-added effect estimates and stated that all of them were significant predictors of student achievement under a random assignment. The teacher value-added estimates had better prediction accuracy when prior student test scores and classroom characteristics were controlled in the models together. Chetty (2013) confirmed their findings by testing for bias in value-added measures using previously unobserved parent characteristics and a quasi-experimental design based on changes in teaching staff. Using administrative records of a large urban school district from 1988–1989 to 2008–2009 in grades 3–8, value-added models which controlled for a student's prior test scores were found that can provide unbiased forecasts of teachers' impacts on student achievement.

As an increasing number of states, districts, and schools that adopted value-added models for education purposes, teacher value-added measures played an important role in identifying effective teachers. Many researchers have seen teacher value-added scores as important information to characterize teacher effectiveness (Hanushek & Rivkin, 2010; Nye et al., 2004). Principals rely on teacher value-added to evaluate teacher effectiveness in order to make decisions about teacher hiring or placement (Tennessee Department of Education, 2007; Tennessee State Board of Education, 2011). Policy makers use these data to make guidelines identify effective teachers as well as teacher evaluation reform in order to improve the instruction (Braun, Chudowsky, & Koenig, 2010; Sanders & Horn, 1998).

However, more recently, value-added estimates are controversial in teacher evaluation as measures of teacher effectiveness. One of the criticisms in value-added estimates is the instability of a teacher's effect. In Kimball and others' (2004) study on the relationship between scores on a standards-based teacher evaluation system and student achievement gain measures in a large Western school district, they found that the estimated relationship of the teacher evaluation scores to value-added scores was positive for each grade and subject and for the reading and math composite; but the coefficients were not statistically significant in all cases. For math and reading at grade 3and math at grade 4, the coefficients were not statistically significant. Another recent study by researchers at Mathematica Policy Research has examined the error rates for measuring teacher effectiveness in the upper elementary grades using value-added models applied to student test score gain data. It concluded that the unstable measures of teacher effectiveness caused by the errors might be sufficiently large enough to lead to the misclassification of many teachers in evaluation system (Schochet & Chiang, 2010). These results suggested that there is still much to learn about the validity of value-added measures because the research did not provide strong and consistent correlations between the value-added scores and other various measures of teacher effectiveness.

Student Perceptions of Teaching

Researchers built different models to measure effectiveness process based on teacher actions in classrooms (e.g. Bill & Melinda Gates Foundation, 2012b; Swank, et al., 1989). Measures of teacher processes are complicated and there are various rubrics to evaluate teachers' practices in classroom. Recently, research has found that student views are potentially an important consideration in measuring teacher effectiveness; and a growing number of districts and states use student surveys to gather feedback on classroom teaching and teacher-student relationships as another way to measure teacher process effectiveness (Ostrander, 1995; Peterson, Wahlquist, & Bone, 2000). Student perceptions toward teaching in classroom were important to teaching processes because students are the main source of information about the classroom (Aleamoni, 1981). Follman (1992; 1995) stated that students are the most direct clients of teachers and, thus, have a broader and deeper experience and contact with teachers than others, including principals, administrators, peers, or parents.

Research on the use of student perception surveys in K-12 education has not been extensive; however, studies consistently suggest that student surveys are a reliable measure of teacher effectiveness. Wilkerson and his colleagues (2000) conducted a study of nearly 2,000 K-12 students and found that, comparing with others' perceptions, teacher's self-ratings, principal ratings, and principal summative ratings, student ratings were more highly correlated with student achievement score and were most significantly accurate in predicting student achievement in reading and mathematics. Similarly, Peterson, Wahlquist, and Bone (2000) confirmed that high validity and reliability of using student surveys for teacher evaluation at elementary, middle, and high school levels. Later on, in a study of development and preliminary validation of student perception survey instrument, Balch (2012) found that teacher scores on a student survey have a positive and marginally significant relationship to value-added estimates of teacher effects on student achievement. Further, by using data from a large-scale pilot in Georgia, Balch (2012) also found a strong link between teacher scores and measures of academic student engagement and student self-efficacy. In Raudenbush's (2013) study that focused on large urban districts, their data showed that student rating and teacher value-added scores are highly convergent. Correlations between indicators computed in these two different measurements had a mean of .94 or higher. In a research synthesis work on teacher effectiveness, Goe (2008) summarized that student perception surveys can provide accurate measures of teacher effectiveness, and were worth considering for inclusion in teacher evaluation systems based on consistent findings of the previous research. This continued evidence pointed out that student perception survey rating can provide valid measurement in teacher effectiveness.

Definition of Working Conditions

Schools provide a working environment and professional community for teachers. They make sure that teachers have appropriate teaching assignment; enough access to information, materials and technology; and adequate time to work with colleagues on matters of instruction (Little, 1993). There are alternative ways to conceptualize different aspects of teacher working conditions. In Perie and Baker's (1997) study, working conditions were identified as administrative support and leadership, student behavior and work atmosphere, and teacher control over the working environment. By using a nationwide survey data, Ingersoll (2001) has explored a set of working conditions and their impacts on teacher turnover behaviors. In his study, four particular organizational conditions in schools have been found to be the most important aspects of school organization to teacher retention: (1) the compensation structure for employees; (2) the level of administrative support; (3) the degree of conflict and strife within the organization; and (4) the degree of employee input into and influence over organizational policies. According to Leithwood's framework (2006), working conditions can be divided into classroom level and school level. The classroom level working conditions include the teaching workload, class size, student composition of class, etc. The school level working conditions are school culture (e.g. safety, academic climate, etc.), school structure (school size, location, physical facilities, etc.), and community relations. In an effort to turn around low-performing schools, Futernick (2007) defined working conditions based on "the belief that when given the opportunity to work on a team with other qualified teachers who share the same vision, teachers can actually jolt the school out of its disequilibrium and transform it into a high-achieving school." His working condition elements include: (1) teams, (2) time, (3) physical environment, (4) class size reduction, (5) autonomy and shared governance, (6) leadership, (7) a well-rounded curriculum, (8) external support, and (9) parent/community involvement. More recently, Johnson (2006; 2012) also has conducted extensive literature reviews and explained from case studies in the topic of teacher working conditions. She has noted that working conditions can include (1) physical features such as the suitability of buildings and equipment; (2) organizational structures that influence workload, autonomy, and supervisory and collegial arrangements; (3) sociological components that influence teachers' roles and status as well as experiences with students and peers; (4) political features that define teachers' power and authority; (5) cultural dimensions that frame values, traditions, and norms; (6) psychological issues that may support or diminish teachers personally; and (7) educational policies, such as those related to teacher education, curriculum, and accountability, that may enhance or constrain what and how teachers can teach.

In the past, discussions of teacher working conditions have focused primarily on teacher salaries and benefits, class size, and internal transfer policies, as well as other similar issues addressed in traditional contractual arrangements (Ladd, 2011; Loeb, Darling-Hammond, & Luczak, 2005; Sass, Hannaway, Xu, Figlio, & Feng, 2012). Compared to these typical issues, it is meaningful to examine the most important working conditions, the ones that matter to teacher effectiveness and student outcomes. According to Leithwood (2006) and Johnson (2006), the classroom and instruction related working conditions are shown to be influential to teacher instruction and student learning. The classroom and instruction level working conditions incorporate more direct supports and resources that inform teachers' instruction and meet the needs of students. Several studies have pointed the importance of the classroom and instruction level working conditions to teachers' growth as well as student success. Such support provides teachers the opportunities to learn and improve the instruction; and it also provides teachers a better classroom environment that meets students' needs (Berry, Daughtrey, & Wieder, 2009;

Berry, Smylie, & Fuller, 2008). Leithwood (2006) concluded that working conditions within the classroom matter most, because they relate more to the instruction and teaching. Moreover, Johnson (2012) studied the effects of teachers' working conditions in high-need schools, and found that the specific aspects of the working conditions that matter the most to teachers are not narrowly school level working conditions such as clean and well-maintained facilities or access to modern instructional technology; instead, it is the teaching and instruction related working conditions such as instructional coaching, professional environment, that predominate in predicting teachers' job satisfaction and their students' achievement. However, there is less agreement about how classroom and instruction level support operate.

Working Conditions and Teacher Effectiveness

For a long time, researchers are sought to identify factors that make a difference in teachers' effectiveness. Research indicated that teacher's personal characteristics (e.g. gender, age, etc.) as well as teacher background qualities (e.g. teaching experience, teacher credentials, certification status, etc.) are factors to teacher effectiveness (Clotfelter et al., 2007; Darling-Hammond, 2006; Darling-Hammond & Sykes, 2003). However, observed teacher personal attributes substantially vary across schools; and these differences are weakly correlated to teacher effectiveness (B. Berry, Daughtrey, & Wieder, 2009; B. Berry, Daughtrey, & Wieder, 2010; Betts, Rueben, & Danenberg, 2000; Harris & Sass, 2011). Recently, studies have shifted from examining teacher input to considering effective teaching in the context of where teachers work. Teacher effectiveness is not just about teachers' knowledge, experiences, and other input attributes, but also about the conditions under where they work (Berry et al., 2010; Jackson & Bruegmann, 2009).

A plethora of studies have shown clearly that school as the workplace can enable or constrain effective teachers (Johnson, 1990; McLaughlin & Talbert, 2001; Rosenholtz, 1989). Johnson (1990) has pointed out that teachers are malleable and dynamic within a workplace, which has a professional and supportive environment that encourages their learning and growth. There is research based evidence that improving the conditions of the school as a workplace can increase the teacher effectiveness not only in classroom instruction but also in students' success in learning outcomes. Bryk (2002) claimed that cultivating teacher working conditions in schools are critical factors that are connected to greater teacher effectiveness (Bryk & Schneider, 2002). In Miami-Dade County Public Schools, Loeb and her colleagues (2011) revealed that among teachers, those who are hired to work in more effective schools improved more rapidly over years (Loeb, Kalogrides, & Béteille, 2011).

Researchers have examined the impact of many working conditions, such as school leadership, teacher compensation, school economic status, etc. (Jackson, 2014; Johnson, Kraft, & Papay, 2012; Ladd, 2011; Loeb, Darling-Hammond, & Luczak, 2005; Sass, Hannaway, Xu, Figlio, & Feng, 2012). Others have studied the effects of classroom and instruction level working conditions, such as professional instruction supports offered, time allotted for instruction planning and collaboration (Jackson & Bruegmann, 2009; Jackson, 2014; Johnson et al., 2012; Ladd, 2011; Rosenholtz, 1989; Y. Goddard & Goddard, 2007). In addition to examining each aspect of these various factors, a few researchers also focused on a series of different aspects as well as overall working conditions. In recent studies, Johnson (2012) compared the effects of each specific aspects of teachers' working conditions on teachers' effects in student achievement, and showed that the professional environment, the principal's leadership, teacher collaboration matter more than the school level working conditions such as clean and well-maintained facilities. Similarly, Jackson (2014) examined several dimensions of the teachers' working conditions in predicting teacher effectiveness; and found that average teacher effectiveness is higher, in schools with strong instructional support and in which teachers perceive a high level of collegial support. Many of the relevant aspects of the working conditions might be categorized as a type of classroom and instruction level working conditions within school. The following review of the literature below explored the existing research on various aspects of teachers' working conditions, all of which might be related to instruction practice or classroom management that could influence teachers' effectiveness.

Instructional Practice Support

Instructional practice support includes assessment data, instructional coaching and professional supports that available to teachers to improve instruction and student learning. Teachers' instructional performance is considered a key to student achievement and the most direct way for teachers to deliver knowledge, skills through the instruction in classrooms (Allen, Pianta, Gregory, Mikami, & Lun, 2011; Cosner, 2011). Brophy's (1996) synthesis of research suggested that effective instruction is conducted in a highly supportive classroom environment that is embedded in a caring learning community. In this environment, most of the class time is spent on curriculum-related activities and the class is managed to maintain students' engagement in those activities. Moreover, a longitudinal study of Desimone and her colleagues (2002) examined the effects of professional development on teachers' instruction. They found that professional development focused on specific instructional practices increases teachers' use of those practices in the classroom. In a recent study of Thomas & Green (2015), they showed that the instruction supports, including instructional practice strategies and instructional professional development, would allow teachers to better establish and communicate clear learning goals to students, monitor their progress, providing feedback, encouraging cooperative learning. These could be helpful in improving students' ability to understand and use knowledge.

Researchers also pointed out that the more information teachers know about their students' learning in classroom, the better they could help in enhancing student performance (Blanc et al., 2010; Datnow, Park, & Kennedy-Lewis, 2012; Hamilton et al., 2009; Tyler, 2011). The assessment data can provide information on what students know, what they should know, and what can be done to meet their academic needs, which may help teachers identify areas of the curriculum that their students need to review, and provide guidance for instructional planning (Blanc et al., 2010; Datnow, Park, & Kennedy-Lewis, 2012; Hamilton et al., 2009). In two qualitative studies of Blanc (2010) and Hamilton (2009), they found that teachers may know more about their students learning progress through the assessment data as supports; and hence can make informed decisions about how to improve student achievement. In one empirical study, Tyler (2011) examined the extent to which a teacher's use of data tools is linked to his/her instructional practices and student achievement. Based on using the data in one mid-size urban district from the 2008-2009 and 2009-2010 school years, he found relatively low levels of teacher interaction with use of assessment data that could potentially inform teaching practice. And he also found no evidence that teacher usage of student data is related to student achievement, but there is a reason to believe these estimates are downwardly biased. While in a more recent empirical study that draws on data from New York City public elementary schools, Jackson (2014) found that average teacher effectiveness measured by student achievement gains is higher, on average, in schools with strong data use.

In addition, opportunities in instructional coaching and professional development supports for teachers have long been seen as a key mechanism for improving teacher effectiveness in classroom instruction and student achievement (Ball & Cohen, 1999; Borko, 2004; Elmore & Burney., 1997; J.W. Little, 1993). It involves teachers to enhance and acquire knowledge and skills in a supportive learning environment (Weiss, et al. 2004). Recent studies confirmed the positive findings on the link between instructional supports and teacher effectiveness for student achievement. In a mixed method study combining both quantitative data from online questionnaires and qualitative data from personal and telephone interviews, Stachler et al. (2013) suggested that collaborative, extended professional development is sustainable and effective of integrating science content into agricultural education curricula, and hence to promote student course achievement. To focus on the student achievement gains of teachers, Biancarosa, et al. (2010) conducted a longitudinal study of a hierarchical value-added model to compare student literacy learning over three years. They found that the use of school-based literacy coaching as an instructional coaching support increasingly enhanced student literacy learning during the implementation of the coaching program. Allen, et al. (2011)

also found substantial gains in student achievement in the year following the intervention of the teacher-student interactions in the classroom via a series of professional development activities that related to instruction practice (e.g. workshop training, personalized instructional coaching).

Although previous studies have consistently shown the importance of the use of assessment data, instructional coaching and professional supports in classroom instruction and students' learning achievement, literature focused on examining the effect of these instruction supports as working condition is still inadequate. The research to date on instruction practice and support has mostly been qualitative and focused on how it affected the instructional practices and teaching progress. There is very limited research that has provided consistent empirical evidence about the impact of instruction practice and support provided by schools on effective teaching and student achievement.

Teaching Workload

Teaching workload refers to the available time for teachers to plan, provide instruction, and eliminate barriers to maximize instructional time during the school day (Ladd, 2011). Teachers view protection of their instructional time as an important working condition for them to support effective teaching in schools. Workload issue about instruction is a continuing problem for teachers. Ingersoll (2003) found that teachers have little time to plan, collaborate, or be involved in curriculum or instructional decisions. Teachers may spend a significant amount of time in doing other non-instruction of work, such as routine paperwork or other duties. The Quality Learning and Teaching Environment (OECD, 2009) survey administered in ten school districts in Georgia found that teachers do not have time during the school day to collaborate and discuss instruction plans with their colleagues. The needs of teachers for time to focus on the instruction duties and the need for the school as an organization to involve teachers in school-time activities outside the classroom create tension for teachers and schools (Adelman, Eagle, & Hargreaves, 1997).

Louis (1996) found that common planning time for teachers is a key to encouraging curriculum innovation, as well as teachers' effectiveness in teaching students. Similarly, other investigators have pointed out that more time for planning and collaboration were correlated highly with teachers' plans to remain in teaching. These workload factors were also related to improved student achievement (Berry & Fuller, 2007). As Rice (2009) pointed out, allocation of enough workload related to teaching can benefit teachers in common planning and collaboration, and hence is intended to support teachers' improvement by learning from their colleagues.

Instruction Resources

According to Ladd (2011), instruction resources refer to the availability of instructional materials, technology, communication, and other instruction related resources to teachers. A teacher needs the necessary equipment, supplies, and materials to work and teach effectively. It is important that teachers have adequate working resources, instructional materials, and up-to-date technology that is easily accessible (Berry, et al., 2008; Loeb, et al., 2004). Studies have linked school quality in resources to teacher performance and satisfaction (Buckley, et al., 2004; Johnson et al., 2012). The findings indicated that poor conditions, especially in school facility and resources, make it more difficult for teachers to deliver adequate instruction to their students.

Some studies have examined the relationship between school resources/facilities and student achievement (Chan & Petrie, 2000; O'Neil & Oates, 2001). In a synthesis study reviewing 141 published studies, Schneider (2002) concluded that almost all of the studies have found statistically significant correlations between the overall working resource and student achievement. In general, students attending school in better environment, more adequate resources had scores of 17 points higher on standardized tests than those attending schools with less resource. More particularly, researchers discussed the need for working and teaching resources that promote students' academic learning through the planned use of space, and found these conditions had impact student achievement as well (Picus, et al., 2005).

School resources not only influence student achievement, but can also influence the work and effectiveness of a teacher. Earthman (2002) found that poor school resources, such as those with poor instructional technology quality, poor maintenance, decreased effective instruction and teaching. In a study focused on reporting the ratings of teachers' perceptions of school resource conditions affected their job performance and teaching effectiveness. Schneider (2003) found that a significant number of teachers noted that poor communication technology quality and inadequate instruction materials, it affected teachers' satisfaction to teach. More recently, by using data from a national survey of school principals in 2005, Duyar's (2010) work empirically investigated the relationship between school facility and resources and the delivery of instruction by controlling school and student characteristics. The findings indicated that the school facility and resource conditions, including air conditioning, physical condition of buildings, and instruction technologies, internet connection are statistically and positively associated with the delivery of instruction. And these facility and resource conditions accounted for 43 percent of the explained variation in the delivery of instruction, a medium-sized effect.

Although some prior research indicated that school resources/facilities were related to student achievement (Chan & Petrie, 2000; O'Neil & Oates, 2001), the findings on the impact of instructional resources on achievement were not consistent. Some studies found weak or no effects of instruction resources on students' achievement (Johnson, Kraft, & Papay, 2012), while some others have shown the opposite effect (Gerber, et al., 2001). For example, Johnson, Kraft, & Papay (2012) found a weak and nonsignificant relationship between teachers' perceptions of the school resources and student achievement when considering a broad range of working conditions; the authors noted that the extent to which teachers have access to sufficient instructional materials and technology are less important because it does not directly relate to the social context of teaching and learning, which heavily influenced teachers' work in a collaborative environment. Moreover, in an empirical study using rigorous quantitative methods (i.e. hieratical linear models), Gerber et al. (2001) indicated that some personnel instruction resources, such as teacher aids, did not have a significant impact on students' achievement in K-3 math or literacy. Particularly, they found a negative effect of personnel resources on student achievement when teacher aid engaged in a range of tasks during instruction (e.g. computer use, small group study, etc.). Due to inconsistent findings it is somewhat unclear which instructional resources are more valuable to teachers.

Classroom Autonomy

Classroom autonomy in classroom refers to how much teachers have control or role in their classrooms over multiple areas of planning and teaching activities. Classroom autonomy is a complex aspect of teachers' working conditions because it requires that educators balance the need for cohesion and structure in school systems against the need for independence in instruction (Campbell 2006; Firestone 2001; Ingersoll 2006). Researchers have argued that teachers require some degree of autonomy to use their professional judgment to tailor instruction to students in various situations and contexts (Glazer 2008). However, research suggests that some limits to autonomy may be necessary, as school administrators and policymakers must consider local and national expectations of accountability, standardization, and equity (Finnigan and Gross 2007; Hanushek and Raymond 2004).

Research pointed out that having more control over classroom decisions, such as selecting curriculum, class assignment and designing discipline is important to teachers. Research found that classroom autonomy is positively associated with teachers' job satisfaction and teacher retention (Guarino, Santibañez, and Daley 2006; Ingersoll & May 2012). Teachers who perceive that they have less autonomy are more likely to leave their positions, either by moving from one school to another or leaving the profession altogether (Berry, Smylie, and Fuller 2008; Boyd, Lankford, Loeb, and Wyckoff 2008; Ingersoll 2006; Ingersoll and May 2012). Classroom autonomy is an important topic for administrators and policymakers to consider when trying to improve teacher satisfaction and reduce teacher attrition rates. A recent study showed that the lack of control over classroom decisions was cited as a primary reason teachers leave the classroom

(Reichardt, et al., 2008). When one group of teachers in a low-retention school was asked why they leave, they responded, "Most of the curriculums are prescripted. We have to follow them and there is no time to be creative. This leaves no space to address students' needs; we have difficult kids and sometimes need to be creative to meet their needs." (p.12). An empirical study by Berry (2007) on National Board Certified Teachers (NBCTs) suggested that over half of teachers were more likely to teach in a school where they can have control over classroom decisions, such as selecting curriculum and designing discipline policy, rather than a school with significantly higher salaries. The financial incentives alone will not attract teachers to high-needs schools.

Classroom autonomy in classroom allows teachers to provide more appropriate help to the various needs of students (Leithwood, 2006; Little, 1990). Teachers seek the flexibility needed to shape their teaching for the diverse learners in their classrooms (Firestone & Pennell, 1993). However, the research to date on classroom autonomy has mostly been qualitative and focused on its impact on teachers' satisfaction or teacher retention. Little empirical work has specifically investigated the impacts of classroom autonomy on improving effective teaching and student achievement gains.

Student Conduct Management

Managing student conduct is one of the school policies and practices designed to address student conduct issues that ensure a safe classroom environment for teachers. Student behavior management is one of the contemporary issues frequently being faced by the schools and affects the workplace for teachers. According to a study designed to understand high school working conditions from the teacher's perspective, they found that student conduct management is one of the most salient aspects of the workplace (McLaughlin & Talbert, 2001).

Several early empirical studies have found that management behaviors, which include managing disruptive behavior among learners, were statistical significantly correlated with student achievement (Oday, 1984; Short, 1987). McGarity & Butts (1984) pointed out that in addition to the student achievement, how schools manage student conduct issues also have a positive impact on student satisfaction and engagement in learning. A recent study confirmed the positive correlations between student behaviors and student achievement; and also concluded the important role of school administration in managing students' behavior through policies and procedures (Nooruddin & Baig, 2014). When there are clear policies and procedures for managing student conduct, teachers are likely to be more satisfied and more effective in classroom.

Moreover, researchers pointed out that the ability to manage student conduct is considered important to teacher effectiveness, because it has a definite effect on teachers' instruction and safety at schools (McKinney, et al., 2005; Obenchain & Taylor, 2005). Researchers have argued that student behavior problems such as bullying, violence in school and other misconduct issues lead to a poor environment for the school community and a sense of fear and frustration in the school culture (Kendziora & Osher, 2009; Liu & Meyer, 2005). Charles (2008) confirmed their conclusions and found that disruptive student behavior can negatively affect and interfere with the ability of a teacher to teach as well as students to learn in a school. Nevertheless, little empirical work has specifically investigated the effects of student conduct management on teacher effectiveness. More research is needed to better understand the effect of student conduct management on student perceptions of teaching and student achievement gains.

Differences in These Factors between High-need and Low-need Schools

The high-need schools serve higher proportion of students at risk of educational failure or otherwise in need of special assistance and support, such as students who are minority, come from low-income families, who are far below grade level, who are at risk of not graduating with a diploma on time, who are homeless, who are in foster care, who have been incarcerated, who have disabilities, or who are English learners (as defined in the Race to the Top application). Retaining effectiveness teachers for high-needs schools may be the most vexing problem facing America's education policy makers. Recently, many studies have consistently shown that inequality in the distribution of effective teachers among schools is to the clear disadvantage of the minority or low-income students (Clotfelter, Ladd, & Vigdor, 2011; Lankford, Loeb, & Wyckoff, 2002). An early report recommended by Education Trust (Peske & Haycock, 2006) indicated that large differences existed between the qualifications of teachers in the high-need schools and low-need schools in Ohio, Illinois, and Wisconsin. Although the gap between the qualifications of teachers in high-need schools and low-need schools has narrowed since 2000, Boyed and his colleagues (2008) still found significant differences in various teacher qualifications, such as teaching experience, student achievement gain, teacher certification, etc. by using data from New York Department of Education. In another study of using statewide administrative data sets from North Carolina and Florida, Sass and his colleagues (2012) found that student achievement gains in elementary school teachers' performance are much higher in low-need schools than in high-need schools.

Recent large-scale quantitative studies provided further evidence that poor work environments and conditions that matter to teachers are more common in high-need schools that attended by minority, limited English proficiency or low-income students (Berry, Smylie, & Fuller, 2008; Johnson, Kraft, & Papay, 2012; Ladd, 2011). In a study using datasets from National Center for Education Statistics' (NCES) surveys, Schools and Staffing Survey (SASS) and its supplement, the Teacher Follow-up Survey (TFS), Ingersoll (2001) found that teachers who left jobs in high-need schools often cite lack of resources, intrusions on instructional time, inadequate time to prepare, and student discipline problems as reasons for quitting. The study claimed that working conditions in high-need schools are often worse than other types of schools, and such conditions are a major cause of high teacher turnover in many schools. More recently, in Ladd's (2011) study that focused on various aspects of working conditions drawing on the North Carolina Working Condition Survey data, teachers' perceptions of working conditions were found to vary by school context (eg. high poverty schools, high minority schools). Similarly, in Jonson, Kraft and Papay' (2012) study of the schools' teaching conditions in Massachusetts, they found that 53 percent of teachers in the low-need schools strongly agreed that their school is a good place to work and learn, compared with just 32 percent of teachers in the high-need schools. Particularly, the specific aspect of working conditions, community support, resources and facilities, and professional expertise, were statistically poorly perceived by teachers in high-need schools that serve a large amount of minority and low income students.

Summary/Literature Gap

Despite growing recognition of the importance of working conditions to teacher effectiveness, researchers have only begun to understand the importance of the classroom and instruction level working conditions on teachers' effectiveness. In most studies, working conditions were either operationalized as a single variable or multiple aspects of working conditions were analyzed (Jackson, 2014; Johnson et al., 2012; Ladd, 2011); few studies have focused on the overall and detailed aspect of classroom and instruction level working conditions. Several earlier researchers have focused on investigation of the effect of classroom-instruction level working conditions such as instructional professional development, teaching workload, etc. on teacher effectiveness (Cohen & Hill, 2001; Jackson, 2014; Rosenholtz, 1989; Wei, Darling-Hammond, Andree, Richardson, & Orphanos, 2009); but these studies did not examine the overall effect and compare the difference in each aspect's effects simultaneously. There is a need to get a deeper understanding of the classroom-instruction level working conditions, and fill the gap in understanding how they differ in affecting teacher effectiveness. Besides, plenty of studies have focused on examining the effects of traditional aspects of working conditions, such as salaries, school leadership, etc. There is a lack of empirical evidence to support the importance of teachers' perceived working conditions in instructional supports and classroom management to effective teaching.

In addition, the research focused on multiple measures of teacher effectiveness that incorporated teacher background quality, student perception about teaching and teacher value-added scores is limited. Many studies were more likely to simply use teacher quality as a measure of teacher effectiveness because of data limitation. Even more studies were restricted to whole-school achievement measures and rarely used controls for past test performance and student background. Instead of teacher value-added measures or student perceptions in teaching, the dependent variables have often been achievement levels. The use of students' perceptions about classroom teaching as a measure of teacher effectiveness is also limited. Besides, most literature focused on the teacher effectiveness in classroom teaching is based on small sample sizes, using qualitative methodologies, and thus, has limited generalizability. In a synthesis study on how working conditions matter to teacher effectiveness, Berry (2010) concluded that more finely tuned research still needs to be conducted to gauge the most critical working conditions linked to effective teaching and student achievement gains.

For the data source, most empirical work has investigated the impact of teacher perceived working conditions on teacher effectiveness by using state administrated data (Johnson et al., 2012; Ladd, 2011; Loeb, Darling-Hammond, & Luczak, 2005); and few has incorporated the nationwide data sets across different states.

Lastly, since more recent research has consistently shown that the inequality in distribution of effective teachers among schools is to the clear disadvantage of high-need schools (Clotfelter, Ladd, & Vigdor, 2011; Lankford, Loeb, & Wyckoff, 2002); and previous research has found empirical evidence that teachers' effectiveness is much stronger in low-need schools than in high-need schools (Boyed, et al., 2008; Sass, et al., 2012). However, in exploring the teacher effectiveness gap, most studies only investigated the difference in teacher characteristics and background qualities, but did not investigate how the working conditions contribute to the teacher effectiveness gap. The research focused on disparities in the overall working conditions and specific aspects of working conditions that affect teacher effectiveness among schools is limited. Not all aspects of working conditions were considered by previous researchers, such as the instructional supports, or student conduct management.

Therefore, it is necessary to conduct more comprehensive research which takes into consideration both teacher value-added scores and student perceptions of teaching as multiple measures of teacher effectiveness, and how instruction-classroom level working condition, as well as each aspect of working conditions, contribute to the variability in teacher effectiveness, particularly between high-need and low-need schools. Advancing our understanding of these relationships is particularly important, as this can give rich information about which working condition factors and school characteristics warrant special attention to recruit, retain and develop effective teachers.

CHAPTER 3

METHODOLOGY

This chapter outlines the methods of the study. It presents a description of the data sets, survey instrument, and items in the data set. It also presents an overview of the analysis plan and the statistical procedures used. In this study, the relationships among teacher value-added score, student perception about teaching, teacher background quality, working conditions, and school context are tested. There are two main domains that are included in the model. The first one is teacher effectiveness constructs (i.e. teacher value-added scores in math and English Language Arts (ELA) and student perception about teaching) and teacher background quality. And the second one is classroom-instruction level working conditions, including managing student conduct, instructional practice, teaching workload, instructional resource, and classroom autonomy.

Research Questions

A conceptual model of working conditions and teacher effectiveness was developed based on previous research and theory. More specifically, the conceptual models included how teachers' working conditions affect teacher value-added scores and student perceptions of teaching, controlling for teacher background qualities. To better understand these issues, this study aimed to assess the relationships of a broad range of teachers' working conditions as well as measures of teacher effectiveness. In addition to the overall classroom-instruction level working conditions, the effects of each working condition were separately examined. The study tested and compared the extent to which of each working condition contributes to the variation in teacher value-added score and student perceptions of teaching. It was hypothesized that students' evaluation of teaching is a mediator between certain working conditions and teacher effectiveness in increasing students' average gain scores. Furthermore, it examined whether all constructs and relationships vary across high-need school context. The study aimed to address the following specific research questions:

1) What are the relationships among teacher value-added score, student perceptions about teaching, and teacher background quality in elementary and middle public schools?
2) What is the effect of each of the working conditions (i.e. instructional practice support, teaching workload, instructional resource, classroom autonomy, and support for managing student conduct) on teacher value-added score and student perceptions, and which work conditions matter more in elementary and middle public schools?
3) How do the impacts of working conditions on teacher value-added score and student perceptions differ according to high-need school context (i.e. minority, free and reduced lunch)?

Data Sources

All data in this study were drawn from the Teacher Working Condition Survey (TWCS) and Student Perception Survey (SPS) and teachers' matched student achievement scores in the Measures of Effective Teaching (MET) Project. The MET project was one of the largest education studies conducted in the United States. Researchers from the University of Michigan helped in collecting a variety of indicators of teaching quality focused on fourth to ninth grade over a two-year period, from the year 2009-2010 to the year 2010-2011. A total of 2741 teachers in 317 schools took part in the MET Study in year one, and 2086 teachers in 310 schools remain in year two of the study. All teachers were located in six large school districts. These included Charlotte-Mecklenburg (NC) Schools, Dallas (TX) Independent School District, Denver (CO) Public Schools, Hillsborough County (FL) Public Schools, Memphis (TN) City Schools, and the New York City (NY) Department of Education in the United States. The schools and teachers that were recruited followed a process of opportunity sampling. In this process, the primary sampling units, districts, were selected as a matter of convenience by the MET Study. The schools within these districts were volunteers that met certain restrictions, and the teachers within schools were volunteers as well.

The measures of students' achievement scores in each teacher's classroom were drawn from state-administered assessments and supplemental achievement tests. At grades 4-8, student learning was measured by state assessments in Math and English Language Arts (ELA) separately; and the ACT Quality Core "end-of-course" assessments for Algebra I, English 9, and Biology were administrated for grade 9 students. For this study, grades 4-8 data on Math and ELA was included.

Student Perception Survey

In addition, students in the sample teachers' classes responded to the student perceptions survey in both years. This survey had about 80 items asking students about their perceptions and attitudes toward their teachers and teaching, which was originally developed by Harvard researcher Ron Ferguson (Bill & Melinda Gates Foundation, 2012a). The survey instrument assessed the extent to which students experience the classroom environment as engaging, demanding, and supportive of their intellectual growth. It includes seven constructs: caring, captivating, conferring, controlling, clarifying, challenging, and consolidating. Care measures students' perceptions of whether the teachers care about them in classroom. Control measures students' perceptions of management of student behaviors in classrooms. Clarify measures students' perceptions of teacher behaviors that help students' to better understand the content being taught. Challenge measures students' perceptions of classroom rigor and required effort. Captivate measures students' perceptions of how well the teacher captures the attention and interest of students. Confer measures students' perceptions of how much a teacher takes students' points of view into account when teaching. Consolidate measures students' perceptions of how much the teacher helps students cognitively represent what they have learned in a connected way and how well the teacher promotes students' understanding of the interconnectedness of different curriculum topics.

The Tripod Survey Assessments is comprised of evidence-based questions and has been shown to reliably predict student achievement gains (Bill & Melinda Gates Foundation, 2010a). The study from Hanover Research (2013) has confirmed that the Tripod Survey can accurately predict student achievement gains and ultimately found that the only thing better at predicting a teacher's test-score gains was previous test-score gains. Moreover, in a more recent study of using MET data, Polikoff (2015) also examined the reliability and validity of the student survey measures. The results showed that the student survey subscales are highly internally consistent (alpha > .80). This study also concluded that the student survey measures of teaching effectiveness are consistent with and even more stable than the value-added estimates of teacher effects on student achievement. This continued evidence pointed out that the Tripod Survey Assessments which will be used in this study can be a valid measurement of teacher effectiveness.

Teacher Working Condition Survey

The Teacher Working Condition Survey were administered in year one of the study (2009-10) to all participating MET teachers. This survey had more than 200 items asking teachers to report on many different features of their school, which include many aspects varied from managing student conduct, instructional practice, teaching workload, instructional resource, classroom autonomy, community support, to school leadership. The TWCS instrument was originally developed by the North Carolina Professional Teaching Standards Commission who completed a literature review of the role of working conditions on teacher satisfaction and teacher mobility (2014 North Carolina Teacher Working Conditions Survey, 2014). The work was spurred by state and national survey data from the National Center for Education Statistics' School and Staffing Survey and focused on teacher identified areas and conditions that drove their satisfaction and employment decisions. Areas identified by teachers included administrative support, autonomy in making decisions, school safety, class size, and time. Survey items were intended to measure aspects of school policies and procedures, supports for technology, professional development and learning, school improvement processes and planning, teacher participation in decision making, school personnel practices, and teachers' beliefs about various aspects of teaching and learning.

Survey validity and reliability were established through prior use and factor analysis of the studies. The reliability testing for the North Carolina Teacher Working Conditions Survey (NCTWC) confirms that the survey is generalizable and will produce similar results with similar populations. The reliability analyses for the NCTWC produced Cronbach's alpha coefficients ranging from 0.86 to 0.96 (2014 North Carolina Teacher Working Conditions Survey, 2014). As part of the MET Project, the Swanlund's (2011) work examined the NCTWC by analyzing data from 286,835 educators from 11 states across the U.S. He concluded that the survey offers a robust and statistical approach for measuring teachers' working conditions, and is capable of producing consistent results across participant groups. Similarly, Clifford et al. (2012) also examined the survey for reliability. They reported the reliability coefficients ranged from .80 to .98 among subscales and with an average of .91 for total. Concerning validity, Clifford et al. (2012) reported that the NCTWC's content validity was "established through an extensive literature review, item measure correlations, and the fit of items to model expectations" (p) and that validity was further established via Rasch analysis.

Sample

MET data is appropriate for this study; it provides detailed information on teacher's working conditions, including managing student conduct, community support, instruction practice and support, teaching workload, and school resources. These factors are related to school-based accountability and teacher effectiveness. Furthermore, MET data provides various measures of teacher effectiveness that including student achievement gains, student perceptions toward teaching, and teacher background qualities. This large scale data focused on elementary and middle school teachers would allow us to explore the relationship of school work conditions to teacher effectiveness, by modeling the relationship between working conditions and teacher effectiveness constructs.

For this research, data were pooled across grades and districts; the two content areas, Math and ELA, were analyzed separately. In the academic year 2009-2010, students in the sample teachers' classes responded to the student survey using either paper or online administration, as chosen by each participating school. And students' achievement scores in state assessments were administrated from March to May in 2010; as well as the same period in the baseline year of 2009. The sampled teachers responded to the Working Condition Survey from July to November in 2009.

In this study, a total of 2026 teachers in 232 schools from fourth to eighth grade in the academic year 2009-2010 were selected. The dataset in year one of MET study was selected because the Teacher Working Condition Survey was only administrated in 2009. And because this research is focused on US elementary and middle school teachers, we only used the 4th-8th grade data. The full sample included four sections of multiple data sources administrated at the same school year: (1) Student achievement gain scores (Teacher value-added score) in Math and ELA in year 2009-2010 and baseline year 2008-2009; (2) Students' responses to Student Perception Survey in year 2009-2010; (3) Teachers' background qualities from the administration data; and (4) Teachers' responses to Teacher Working Condition Survey in year 2009-2010.

Table 3.1 *Teacher Characteristics*

Teacher Characteristics	Percentage
Gender	
Female	83.5%
Male	16.5%
Race/ethnicity	
White	56.9%

Black	35.2%
Hispanic	5.7%
Other	2.2%
Teaching subject	
Math	63.6%
ELA	68.9%
Teaching grade	
4th	21.5%
5th	21.6%
6th	22.0%
7th	18.7%
8th	16.2%
Teacher background quality	
Master or higher degree	36.2%
Teaching Experience	
First Year	3.4%
2-3 Years	16.4%
4-6 Years	21.1%
7-10 Years	22.1%
11-20 Years	25.6%
20+ Years	11.5%

In Table 3.1 we can see the distribution of the sample teachers. Our sample was constructed predominantly of female teachers who identified as White or Black. The sample teachers were almost equally distributed across grades 4 through 8 and between Math and ELA. We can see an overlap percentage in Math and ELA subjects because of the possibility that elementary teachers usually taught in both subjects. Besides, in the sample there was about 36 percent of teachers who have advanced educational degrees, and the majority was junior career teachers (teachers whose teaching experience is 5 to 20 years).

Table 3.2 *School Context*

School Context	Mean	S.D.
Student race/ethnicity composition		
White	.236	.236
Black	.366	.317
Hispanic	.311	.232
Other races		
Proportion of students in special programs		

Gifted Program	.074	.134
English Language Learner Program	.130	.118
Free and Reduced Meal Program	.680	.246

In addition, Table 3.2 showed the descriptive statistics of the school context. In the sample schools, averages of around 10 percent of students participated in the gifted program, or the English language learner program. Meanwhile, the sample schools had a large proportion (>50%) of students who are identified as minority (i.e. non-white) or poverty (i.e. Free and Reduced Meal Program), on average. This could be the reason that more disadvantaged schools were more likely to be volunteered to participate in the MET project to receive a bonus or extra funding.

Data Preparation

Data cleaning was done through SAS program. All dependent and independent variables were prepared at teacher level. The students' responses from student survey were aggregated into teacher level by calculating the average ratings. Multiple working conditions, teacher value-added score, the average student perceptions, and teacher qualities were merged through the unique teacher ID in the MET study.

The 2026 cases were examined for missing values and accuracy. Survey items selected for the teacher working conditions asked participants to describe their extent of agreement with each item from strongly disagree to strongly agree along a 4-point Likert scale. Participant responses in the survey of 'don't know' had been previously coded with a value of 5 on a 4-point scale by the MET researchers. As a result, 'don't know' responses were recoded to 'missing' to avoid skewing the numeric value that would be later computed from the continuous scale. This data cleaning did not change the sample size of 2026 cases.

In the cases dealing with any missing data values in the following model analysis, Multiple Imputation was utilized. The Multiple Imputation (MI) approach prevents large further reduction of sample size and preserves unbiased estimates of population parameters by creating multiple data sets that replace each missing value with two or more plausible values (Acock, 2005; Rubin, 2004). Compared to more traditional approaches used to deal with missing data, Multiple Imputation also preserves statistical power and achieves more stable estimates. Particularly, Multiple Imputation is more appropriate and has better statistical properties than other strategies for SEMs (Allison, 2003). After using the Listwise method to delete all missing values, it resulted in a significant reduction of the sample size from 2026 to 1450 cases. It is necessary to retain enough sample size to assure the statistical power for such complex structural equation models analysis. By using SPSS statistical software, missing data patterns were analyzed and five new data sets were created using data imputation and augmentation process. New values were inserted into each augmented data set that represented the uncertainty about the right value to impute (Yuan, 2010). After comparing the descriptive statistics of all variables that were included in the imputation process in each augmented data set, most appropriate augmented data set were selected and prepared for further SEM analysis. The imputed sample and original sample had similar descriptive statistics for all variables.

Measures

Teacher Value-added Score

MET researchers used student test scores to construct "value-added" measures of teaching effectiveness for individual teachers (White & Rowan, 2014). Firstly, the MET researchers created a roster to track every student in classroom for each teacher. Then by using these rosters, the value-added measures were estimated by connecting to roster data of student achievement and their teachers in class. In addition, MET researchers estimated value-added models based on a single outcome measure separately by the state mathematics and English Language Arts (ELA) test. A context adjusted 2-level hierarchical model was used to estimate teacher value-added scores (Raudenbush & Bryk, 2002):

Level 1 (student level): $A_{ijt} = \beta_{0j} + \beta_{1j} PA_{i(t-1)} + \beta_{2j} X_{ij} + r_{ij}$

Level 2 (teacher level): $\beta_{0j} = \gamma_{00} + \gamma_{01} \overline{PA}_{0jk(t-1)} + \gamma_{02} \bar{X}_{0jk} + u_{0j}$

where: (1) A_{ijt} is one of the achievement outcome for student *i* taught by teacher *j* in the current school year t; (2) $PA_{i(t-1)}$ is student *i*'s prior achievement test score at year *t*-1; (3) X_{ij} is a vector of student background variables for student *i* taught by teacher *j*; (4) $\overline{PA}_{0jk(t-1)}$ is the mean prior test score of all students in the classroom *k* taught by teacher *j* at year *t*-1; (5) $\bar{X}_{0jk}$ is a vector of the averages of student background variables for the classroom k taught by teacher *j*; (6) r_{ij} is a student-specific random residual error; and (7) u_{0j} is the random residual associated with teacher *j*.

This model accounted for prior achievement in the subject area, student background (e.g. gender, age, ethnicity, free-reduced lunch status, special education, gifted status, etc.), and class-level average prior achievement scores, class-level aggregated student background proportion. The teacher-level residuals were used as estimates of the value-added score for a specific teacher.

Student Perception about Teaching Items

Student evaluation of teaching was measured by students' perceptions toward classroom teaching and learning. Two versions of the survey were administered; one for students in fourth and fifth grades, and one for students in sixth to ninth grades. The versions differed mostly in the wording of questions and scales. Ten items on student perceptions about teaching were selected from the student perception survey in MET, based on theory and operational definitions used in empirical studies. All selected items were identical between the two versions of the student perception survey. Students were asked to indicate their level of agreement on a 5-point Likert scale, ranging from Scale: 1= No never/Totally untrue; 2= Mostly not/Mostly untrue; 3= Maybe/Sometimes; 4= Mostly yes/Mostly true, 5= Yes always/Totally true with the following statements:

1) My teacher in this class makes me feel that s/he really cares about me.
2) My teacher seems to know if something is bothering me.
3) My teacher explains difficult things clearly.
4) My teacher has several good ways to explain each topic that we cover in this class.

5) My teacher knows when the class understands, and when we do not.

6) If you don't understand something, my teacher explains it another way.

7) In this class, my teacher accepts nothing less than our full effort.

8) My teacher checks to make sure we understand what s/he is teaching us.

9) My teacher wants me to explain my answers—why I think what I think.

10) My teacher takes the time to summarize what we learn each day.

Teacher Background Quality Items

There were two measures related to teacher background qualities in the MET data. Other common measures of teacher background quality such as teaching certificates or teacher's license were not included in the MET study. The administrative data provided teachers' education level and teaching experience as two measures of teacher background quality.

Teacher Working Conditions Items

All working condition indicators of each latent variable were selected from the teacher working condition survey in MET, based on theory and operational definitions used in empirical studies. The following sections list items representing different dimensions of working conditions. The response scales for the teacher working condition survey items were four-point Likert-type scales with anchors at: 1= Strongly disagree; 2= Disagree; 3= Agree; 4= Strongly agree.

Instructional Practice Support Items

1) Teachers use assessment data to inform their instruction.
2) Teachers work in professional learning communities to develop and align instructional practices.
3) Provided supports (i.e. instructional coaching, professional learning communities, etc.) translate to improvements in instructional practices by teachers.
4) Teachers are encouraged to try new things to improve instruction.

Teaching Workload Items

1) Teachers have sufficient instructional time to meet the needs of all students.
2) Teachers are allowed to focus on educating students with minimal interruptions.
3) Efforts are made to minimize the amount of routine paperwork teachers are required to do.
4) Teachers are protected from duties that interfere with their essential role of educating students.
5) Class sizes are reasonable such that teachers have the time available to meet the needs of all students.

Instruction Resource Items

1) Teachers have sufficient access to appropriate instructional materials.
2) Teachers have sufficient access to instructional technology, including computers, printers, software and internet access.

3) Teachers have access to reliable communication technology, including phones, faxes and email.

4) The reliability and speed of Internet connections in this school are sufficient to support instructional practices.

Classroom Autonomy Items

1) Teachers are relied upon to make decisions about instruction issues.

2) The faculty has an effective process for making group decisions to solve problems.

3) Teachers are trusted to make sound professional decisions about instruction.

4) Teachers have autonomy to make decisions about instructional delivery (i.e. pacing, materials and pedagogy).

Managing Student Conduct Items

1) Students at this school understand expectations for their conduct.

2) Students at this school follow rules of conduct.

3) School administrators support teachers' efforts to maintain discipline in the classroom.

4) School administrators consistently enforce rules for student conduct.

5) Policies and procedures about student conduct are clearly understood by the faculty.

High-need School Context Items

The high-need school context was considered as school background information in this study. It included the proportion of minority students, free and reduced lunch eligible students. The higher proportions of students in each group indicated the school is more hard-to-staff and in higher-need.

Method: Structural Equation Modeling

The structural equation modeling (SEM) approach was the primary statistical method that was used in this study. SEM is a statistical method that takes a confirmatory approach to data analysis of a structural theory bearing on the relationships of interested variables (Byrne, 1998). SEM collects statistical techniques to examine relationships between multiple independent and dependent variables (Tabachnick & Fidell, 2007), which has proven to be a useful analytical framework for examining complex, interrelated, and multidimensional models (Tomaken & Waller, 2005). Moreover, theoretical constructs that cannot be observed directly, referred to as latent factors, can be assessed with SEM. SEM deals with the relationships between latent variables that are free thou the random error (Loehlin, 1987). In this study, the latent variables were student perception of teaching; instructional practice supports; teaching workload; instruction resources; classroom autonomy, and student conduct management. Assessment of these latent variables was determined by direct measurement of observed variables, thus providing an indirect measure of an underlying construct (Byrne, 1998).

One advantage of using SEM was that it can simultaneously estimate both the measurement structures (Confirmatory Factor Analysis) and the causal relationships (Multiple Regression Analysis) in one full model, which statistically and visually represents the complex relationships between variables (Bollen, 1989). As for a clearer visualization of the relationships between the constructs, SEM presented the standardized estimates and direction of each effect and measurement structures among all of the variables in one graph. Additionally, SEM provides an integrative approach that captures the combined effect of latent variables. Compared to alternative statistical methods (e.g. multiple regression), SEM provides estimates in both direct effects and indirect effects. Beside, SEM analysis can assess the relative substantive validity across models and generate more parsimonious explanations, using parameter estimates and comprehensive fit indices (Kline, 2011). The fit indices and parameter estimates were used to judge the model fit, with acceptable fit statistics supporting the overall latent factor model, significant factor loadings upholding the measurement models, and significant causal parameters supporting the structural model (Schumacker & Lomax, 2010). For all of these reasons, SEM was an appropriate and suitable approach for examining the relationships between the multiple working condition latent factors, student perceptions about teaching, teacher value-added scores, and teacher background qualities.

The SEM analysis was approached in a four-stage process, the first one was testing and refining the measurement models, the second one was evaluating the structural model, the third one was cross-validating the estimated models, and the last one was invariance testing the SEMs across high-need school contexts. The SEM models were created by using the randomly half of the data. The model accuracy and cross-validation was measured by applying the model to the other half of the data.

To control for measurement error, each construct was assessed with multiple indicators, and then a measurement model was estimated. The final measurement models were evaluated based on the parceled indicators for each latent variable. The item parceling method was used to decrease the number of items for each construct

(Schumacker & Lomax, 2010); to enhance the model fit by improving the reliability of the indicators (Little, Cunningham, Shahar, & Widaman, 2002). In this study, item parceling was conducted as a composite indicator comprised of the average of two or more items. Before go to the structural model step, the final measurement models were cross-validated across a second independent sample.

Subsequently, the structural models were estimated and evaluated. The structural model of the latent factors depicted the theorized structural relations among the factors and defined the relations among the latent variables by specifying the manner by which the latent variables directly or indirectly influence changes in certain other latent variables (Byrne, 1998). The hypothesized structural models for Math and ELA teachers were evaluated separately. At last, the estimated structural models were validated by using an invariance testing strategy. Both observed variables and latent variables were included in the structural equation modeling. Observed variables in the structural equation model were teacher value-added score outcomes in math and ELA, and teacher background quality which was assumed to be measured with error. The latent variables, instructional practice supports; teaching workload; instruction resources; classroom autonomy, student conduct management, and student perception were latent constructs, not directly measured but estimated from related indicators.

Data Analysis Procedures

Descriptive Statistics

Since SEM assumes multivariate normality, published recommendations were followed by examining the data for multivariate and univariate normality (Kline, 2011;

Schumacker & Lomax, 2010). Prior to conducting SEM, preliminary descriptive statistics (mean, standard deviation, skewness, and kurtosis of the distribution) and reliability estimates of all indicators were examined. The reliability coefficients of each latent construct and preliminary item analysis were carried out and as a whole and within each dimension of constructs. Correlations among observed indicators were calculated and examined in order to do a preliminary assessment of the correlation pattern. In addition, an exploratory factor analysis (EFA) was used to identify the underlying structure of the items and reduce the number of observed indicators for each scale in this study. The separate measurement models of each scale of the model and full measurement models were specified and estimated. SPSS software was used in analyzing the descriptive and correlation statistics.

Development of the Measurement Models

The measurement model that tests individual model parameters was analyzed using confirmatory factor analysis. Confirmatory factor analysis (CFA) was conducted to test the factor structure of the latent factors. Using CFA, observed variables were tested to determine if their factor loadings were significant and appropriately placed. In addition to the significance level of each parameter, Stevens (2002) also provided a table of critical values against which loadings can be compared. For sample sizes over 1000, he recommended values greater than .162. Items with loadings of .40 or above were considered for inclusion in the subscales. And Cronbach's alpha values of .6 and above for each full scale were considered acceptable.

Once overall fit of the models was established, the construct validity, error variance, indicator reliability, Cronbach's alpha, and construct reliability were assessed. The construct validity was assessed on the standardized factor loadings, which measures the variance that is accounted for by the latent variables. The construct validity intended to see the extent to which indicator converge or shares in a single construct. An indicator has high validity when its factor loading value is high and significant. The indicator reliability was assessed by the square of the standardized factor loadings, which measures the variance in each measured variable explained by the underlying latent variable. Cronbach's alpha was used to measure internal consistency and integrate reliability. The results of the CFAs that yield the factor structure of the indicators and latent constructs were incorporated into the structural model.

Development of the Structural Models

Structural equation models can be developed in five steps (Schumacker & Lomax, 2010). These steps are: 1) model specification, 2) model identification, 3) model estimation, 4) model testing, and 5) model modification. Model specification is the specification of the relationships among latent factors. The causal links were developed based on the conceptual theories and hypothesized models. The structural models were proposed to answer the first three research questions. On the right side of the model (Figure 3.1), the relationships among teacher value-added score, student perception, and teacher background quality were tested that aims to answer the first research question. Various working conditions served as exogenous variables on the left side of the model examined the effects of each aspect of working condition on teacher effectiveness constructs and that aimed to answer the second and third research question. Additionally, in order to address the last research question, an invariance testing of the model was conducted for the multi-group SEM analyses.

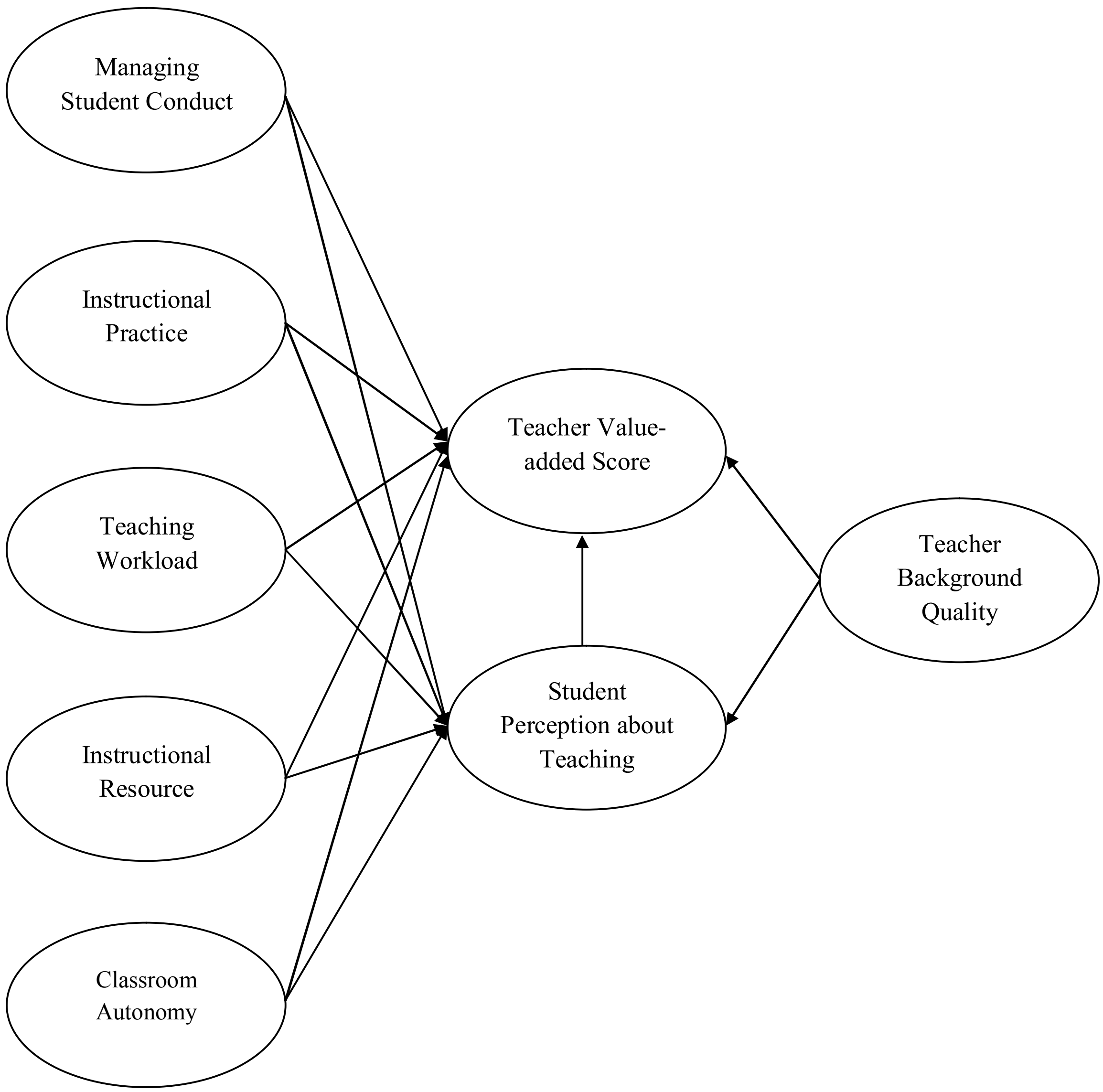

Figure 3.1 Structural Model

Model identification is "to ask whether unique values can be found for the parameters to be estimated in the theoretical model" (Schumacker & Lomax, 2010). This identification process entails examining the total number of parameters that can be estimated, by assessing the model degrees of freedom (*df* = number of observed variances/covariances - number of parameters to be estimated). When the number of observed variances/covariances is larger than the number of parameters to be estimated, the model is overidentified and can be tested for fit.

Once the model was identified, the next step was to estimate model parameters as well as interpret the parameter estimates. The main focus of the estimation process was to yield parameter values such that the discrepancy between the sample covariance matrix and the population covariance matrix implied by the model is minimal. For the analysis of the models, Maximum Likelihood was used as the estimation method.

The next step after estimating model parameters was to evaluate the model fit. To assess the overall goodness-of-fit for the SEM model, a combination of absolute, comparative, and indices of fit were used. According to Schumacker and Lomax's (2010), criteria were used: (a) the chi-square test, the test significance level indicates that whether there is statistically significant difference between the observed and the fitted correlations. (b) Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI) Normed Fit Index (NFI), and Comparative Fit Index (CFI) values close to .95 reflect a good fit and 1.0 indicate a perfect fit. (d) Root-Mean-Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) value less than .05 indicates a good model fit. The RMSEA and SRMR index are adequately sensitive to model misspecification and it is possible to build confidence intervals around these values. These fit indices indicate how well the data supports the theoretical model.

The last step was the model modification. When the initial model did not fit the data well, the model would be respecified and reanalyzed. During this respecification process, possible alternatives in the proposed model due to theory or empirical results need to be considered. Once a satisfactory fit is achieved, the parameter estimates would be examined to determine whether estimates of the parameters were meaningful. This model modification analytic approach in this study followed what Kline (2011) refers to as model generation, the model should be modified with the goal of discovering a model that makes theoretical sense, fits the data, and is relatively parsimonious when the initial model did not fit the data well.

Cross-validating the SEM Models

A cross-validation technique was used to evaluate the predictive accuracy and ensure the validity of the fitted models. Cross-validation is best regarded as a method for establishing the validity of a model by sample splitting and through replication (Bagozzi & Yi, 1988). The purpose of using cross-validation was to establish the model validity by separating estimation from independent samples in order to ascertain that the fit is not a result of idiosyncratic sample characteristics. Cross-validation in SEM is similar to replication; the evaluation of estimated models can be performed through double cross-validation to a different sample (MacCallum, Roznowski, Mar, & Reith, 1994). Cudeck and Browne (1983) have developed a procedure for doing cross-validation within the context of SEMs with latent variables. The cross-validation method used in this study proceeded as follows. First, it began by randomly splitting the sample into two subsamples: a calibration sample and a validation sample. Then the models were fitted or refined based on modifications by using the calibration sample. After that, an invariance testing strategy was used to test for replicability of the full measurement and structural models across calibration and validation samples. If the invariance testing does not reveal a significant difference in estimated parameters between the calibration and validation groups, it suggests that the model validation is successful.

Multi-group Analysis

Before testing the multi-group differences by school context, a Tamhane's T2 test was used to examine the differences in all constructs between high and low need school context. If significant differences in working conditions or teacher effectiveness constructs were found, further SEM invariance testing was used to explore the differences in the relationships between working conditions and teacher effectiveness constructs.

In order to test whether interrelationships among constructs differ across subgroups specified by school context, the invariance testing strategy was used for the multi-group analysis. The procedures for invariance testing using LISREL were guided by Joreskog and Sorbom (1993). The final structural models were used for multi-group analysis as baseline models (Byrne, 1998). In each step, model parameters were then estimated and the fit of the model to the data was assessed based on goodness-of-fit indices, as discussed above. The acceptability of the final structural model was decided based on the results of these model evaluations. If the model was not acceptable, it was further revised based on theoretical credibility and modification indices.

Multi-group analysis for both the measurement model and structural model was conducted using the following logic: First, the baseline models were estimated for each group to identify the best fitting model. Once the baseline models for each group were established, multi-group analyses were performed by comparing the relevant parameters across groups using chi-square difference tests. Second, a joint, unconstrained model was estimated for both groups (i.e., coefficients were allowed to vary freely across groups). Third, a joint constrained model was estimated where the parameters across groups are constrained to be equal to each other. Fourth, the fit of the constrained model was compared with that of an unconstrained model using the difference in chi-square statistic. If the chi-square difference statistic does not reveal a significant (a significance level of .05) difference between the unconstrained and constrained models, then we can conclude that the constrained model applies across groups. If the chi-square statistic reveals a significant difference between the unconstrained and constrained models, then modification indices were used to partially remove the constraints (i.e., to identify where the differences were).

Specifically, the equality of measurement models was first tested. It began with a test of whether the two groups have the same factor structure, which was followed by a test of whether the measures indicate the factors in a different way (the invariance of factor loadings). Given the constrained or partially constrained factor loadings across groups, then whether the structural relationships are invariant could be investigated by using chi-square difference tests. If the chi-square difference between the unconstrained and constrained model was significant, a further investigation continued to point out which structural paths show a significant difference. Modification indices indicate possible structural relationships that can be unconstrained and then improve the model fit of the data.

To sum, this chapter outlined the methods used in the study. It provided an overview of the dataset, sample, items and factors in the study. It laid out the analytic approaches taken to estimate the models.

CHAPTER 4

ANALYSIS AND RESULTS

This chapter presents the results of data analysis and hypothesis testing. The chapter is divided into three sections. The first section outlines descriptive statistics and correlations for all variables of this study. The second section of the chapter presents structural equation modeling as well as the model cross-validation analyses. And the last section presents the multi-group analysis related to whether high-need school context plays a role in these relationships. Both a brief statement about the analyses and the results are provided for each structural model. This part is presented in a brief narrative to allow the reader to follow the analytical decisions made sequentially at each step of the analysis.

Descriptive Statistics

MET is one of the most comprehensive data sources on teachers' effectiveness and students' achievement. It includes students' perceptions, about classroom learning, and teachers' perceptions about working conditions. To understand the data in detail, descriptive statistics for each of the variables, including mean, standard deviation, frequencies and correlation matrices of teacher value-added scores, student perception items, teacher background quality items, working condition items, and school context items are shown below.

Teacher Value-added Score in Math and ELA

The value-added score captures the pure achievement gains for individual teachers by controlling for student background, family SES, etc. The estimated teacher value-added scores for mathematics and English are based on a single outcome measure by the state mathematics and English Language Arts (ELA) test, separately. The ELA teachers (*N*=1395) have an average value-added score of -.0002 with standard deviation .191; and Math teachers (*N*=1288) have an average value-added score of .003 with standard deviation .253. Both of the value-added scores for Math and ELA teachers have acceptable distributions (see Figure 4.1).

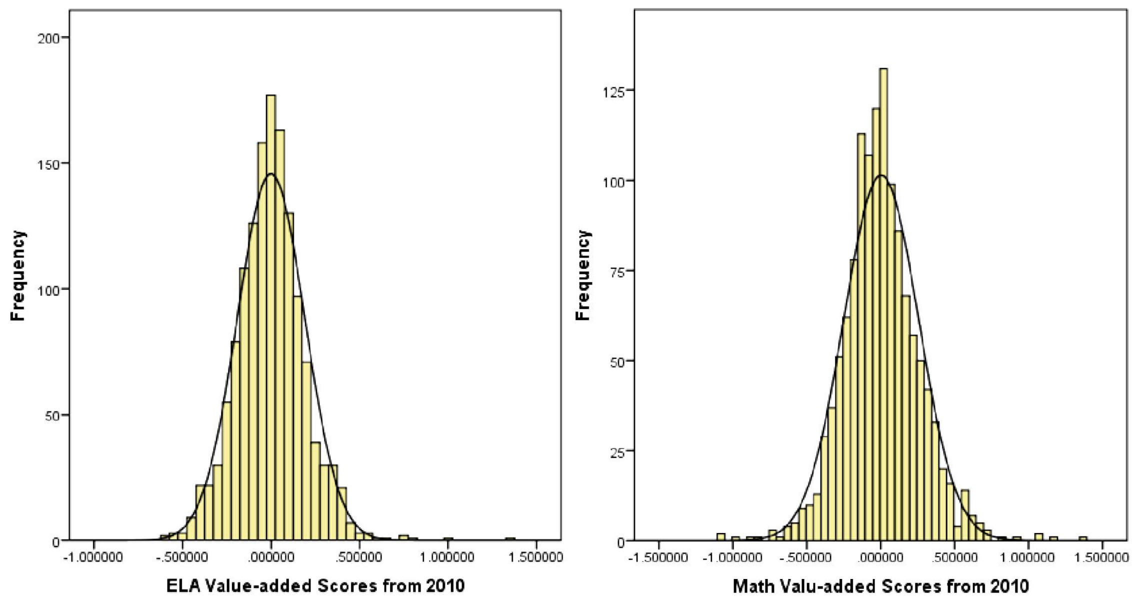

Figure 4.1 Histogram of Teacher Value-added Score

Students' Perceptions of Teaching Items

The student perception survey in MET contained many items about the classroom learning and teaching. There were 10 items that pertain to psychological dimension of students' perceptions about teaching.

Table 4.1 *Descriptive Statistics and Correlations for Student Perception Items*

	Scale and Items	Mean	S.D.	1	2	3	4	5	6	7	8	9	10
1	My teacher in this class makes me feel that he/she truly cares about me.	3.92	.584	1									
2	My teacher knows if something is bothering me.	3.35	.587	.80**	1								
3	My teacher explains difficult things clearly.	4.03	.464	.79**	.74**	1							
4	My teacher has several good ways to explain each topic that we cover in class.	4.04	.437	.81**	.72**	.85**	1						
5	My teacher knows when the class understands, and when we do not.	3.96	.407	.74**	.70**	.79**	.80**	1					
6	If you don't understand something, my teacher explains it another way.	4.16	.409	.75**	.71**	.82**	.85**	.78**	1				
7	In this class, my teacher accepts nothing less than our full effort.	4.11	.403	.54**	.60**	.70**	.65**	.65**	.62**	1			
8	My teacher checks to make sure we understand what s/he is teaching us.	4.21	.438	.77**	.73**	.85**	.85**	.79**	.82**	.71**	1		
9	My teacher wants me to explain my answers—why I think what I think.	4.12	.366	.57**	.51**	.58**	.66**	.61**	.63**	.60**	.66**	1	
10	My teacher takes the time to summarize what we learn each day.	3.61	.497	.64**	.67**	.67**	.72**	.67**	.68**	.55**	.72**	.57**	1

1= No never/Totally untrue; 2= Mostly not/Mostly untrue; 3= Maybe/Sometimes; 4= Mostly yes/Mostly true, 5= Yes always/Totally true

N=1921

Note: **, p<.01; *, p<.05

Since what the teachers do in the classroom is influential in effective teaching, these items were considered as reflecting the teachers' effectiveness in students' view. These are items that capture students' perceptions about classroom instruction and how much they value their teachers and teaching. The Cronbach's alpha reliability estimate of all 10 items was .96. Detailed descriptive statistics and correlation matrix were shown in Table 4.1.

Teacher Background Quality Items

There was no existing teacher background quality composite in MET data. Teachers' education level and teaching experience were used to create a composite measure of teacher background quality. Teaching experience was significantly correlated with Teachers' advanced degree (r=.234). The teacher background quality measure was created by adding these two variables, in a category scale. Measures of teaching certificates or teacher's license were not included in MET.

Table 4.2 *Frequency of Teacher Background Qualities*

Teacher background quality	Percentage
Teaching experience	
1~3 years	19.8%
4~10 years	43.2%
10+ years	37.1%
Advanced education degree	
Have master or higher degree	36.2%
Not have master or higher degree	63.8%
Composite	
1~3 years	16.2%
4~10 years; or 1~3 years with advanced degree	30.8%
10+ years; or 4~10 years with advanced degree	35.3%
10+ years with advanced degree	17.7%

N=1878

Table 4.2 showed the frequency of the teachers' background qualities. In the sample there was about 36 percent of teachers who have advanced educational degrees, and the majority was junior career teachers (teachers whose teaching experience is 4 to 10 years).

Working Conditions

Corresponding to teacher value-added scores and students' perception survey items, teachers were asked to report their perceptions about the school working conditions of their workplace. The working condition items capture various aspects of teacher perceived conditions that may influence the teaching performance of teachers. There were 28 items in total that pertain to teachers' perceptions of working conditions. The Cronbach's alpha reliability estimate of all 28 working condition items was .94. Each of the 5 working conditions items are shown below.

Managing Student Conduct Items

There were 5 items that relate to teachers' perception of support for managing student conduct. These items capture teachers' perception of how school policies and administrator manage student conduct issues. The Cronbach's alpha reliability estimate of all 5 items was .88. Detailed descriptive statistics and correlation matrix were shown in Table 4.3.

Table 4.3 *Descriptive Statistics and Correlations for Managing Student Conduct Items*

	Scale and Items	Mean	S.D.	1	2	3	4	5
1	Students at this school understand expectations for their conduct.	3.04	.805	1				
2	Students at this school follow rules of conduct.	2.55	.845	.60**	1			
3	School administrators support teachers' efforts to maintain	2.79	.869	.55**	.60**	1		

	discipline in the classroom.							
4	School administrators consistently enforce rules for student conduct.	2.56	.944	.57**	.64**	.80**	1	
5	Policies and procedures about student conduct are clearly understood by the faculty.	3.02	.783	.56**	.46**	.55**	.58**	1

1= Strongly disagree; 2= Disagree; 3= Agree; 4= Strongly agree

N=1420

Note: **, $p<.01$; *, $p<.05$

Instructional Practice Support Items

There were 4 items that relate to teachers' perception of instructional practice and support. These items capture teachers' perception of instructional and professional support for improving instruction. The Cronbach's alpha reliability estimate of all 4 items was .80. Detailed descriptive statistics and correlation matrix were shown in Table 4.4.

Table 4.4 *Descriptive Statistics and Correlations for Instructional Practice and Support Items*

	Scale and Items	Mean	S.D.	1	2	3	4
1	Teachers use assessment data to inform their instruction.	3.28	.599	1			
2	Teachers work in professional learning communities to develop and align instructional practices.	3.11	.723	.50**	1		
3	Provided supports (i.e. instructional coaching, professional learning communities, etc.) translate to improvements in instructional practices by teachers.	2.98	.703	.43**	.64**	1	
4	Teachers are encouraged to try new things to improve instruction.	3.19	.667	.42**	.46**	.52**	1

1= Strongly disagree; 2= Disagree; 3= Agree; 4= Strongly agree

N=1345

Note: **, $p<.01$; *, $p<.05$

Teaching Workload Items

There were 5 items that relate to teachers' perception of teaching workload. These items capture teachers' perception of the available time of instruction. The Cronbach's alpha reliability estimate of all 5 items was .76. The item "Class sizes are reasonable such that teachers have the time available to meet the needs of all students." was found poorly correlated to the other items. And considering this item includes the "class size", which may confuse the conceptual meaning of workload with other items, the item was deleted. The Cronbach's alpha increased to .78 for the scale. Although the reliability coefficient of teaching workload construct is lower than that of other working condition constructs, it was above 0.7, which is commonly used as a cut-off value of Cronbach's alpha for a good level of reliability (Lance, 2006; Henson, 2001). After deleting this item for low reliability, we kept 4 items that reflected the teaching workload. Detailed descriptive statistics and correlation matrix were shown in Table 4.5.

Table 4.5 *Descriptive Statistics and Correlations for Teaching Workload Items*

	Scale and Items	Mean	S.D.	1	2	3	4
1	Teachers have sufficient instructional time to meet the needs of all students.	2.56	.772	1			
2	Teachers are allowed to focus on educating students with minimal interruptions.	2.61	.811	.44**	1		
3	Efforts are made to minimize the amount of routine paperwork teachers are required to do.	2.30	.852	.40**	.46**	1	
4	Teachers are protected from duties that interfere with their essential role of educating students.	2.71	.906	.39**	.50**	.51**	1

1= Strongly disagree; 2= Disagree; 3= Agree; 4= Strongly agree

N=1407

Note: **, p<.01; *, p<.05

<u>Instruction Resources Items</u>

There were 4 items that relate to teachers' perception of instruction resources. These items capture teachers' perception in how schools provide adequate working and instruction related resources. The Cronbach's alpha reliability estimate of all 4 items was .81. Detailed descriptive statistics and correlation matrix were shown in Table 4.6.

Table 4.6 *Descriptive Statistics and Correlations for Instruction Resources Items*

	Scale and Items	Mean	S.D.	1	2	3	5
1	Teachers have sufficient access to appropriate instructional materials.	2.98	.788	1			
2	Teachers have sufficient access to instructional technology, including computers, printers, software and internet access.	2.89	.896	.54**	1		
3	Teachers have access to reliable communication technology, including phones, faxes and email.	3.03	.793	.46**	.61**	1	
4	The reliability and speed of Internet connections in this school are sufficient to support instructional practices.	2.92	.840	.42**	.49**	.43**	1

1= Strongly disagree; 2= Disagree; 3= Agree; 4= Strongly agree

N=1451

Note: **, $p<.01$; *, $p<.05$

Classroom Autonomy Items

There were 4 items that relate to teachers' perception about autonomy in classroom. These items capture teachers' perception of how much control they have on their work related activities and in their role as a classroom teacher. The Cronbach's alpha reliability estimate of all 4 items was .75. Detailed descriptive statistics and correlation matrix were shown in Table 4.7.

Table 4.7 *Descriptive Statistics and Correlations for Classroom Autonomy Items*

	Scale and Items	Mean	S.D.	1	2	3	4
1	Teachers are relied upon to make decisions about instruction issues.	2.91	.761	1			
2	The faculty has an effective process for making group decisions to solve problems.	2.73	.810	.67**	1		
3	Teachers are trusted to make sound	2.97	.759	.82**	.64**	1	

	professional decisions about instruction.						
4	Teachers have autonomy to make decisions about instructional delivery (i.e. pacing, materials and pedagogy).	2.77	.877	.47**	.39**	.51**	1

1= Strongly disagree; 2= Disagree; 3= Agree; 4= Strongly agree

N=1392

Note: **, p<.01; *, p<.05

High-need School Context Items

There were two items that related to the high-need school context: the proportion of minority students, and students living in poverty. The higher proportions of students in each group indicated more hard to staff and higher-need school context. The minority proportions are those of non-white students. The poor students are those eligible for the free and reduced meals (FARM) program. Since only 167 out of 232 schools have reported their school context information in the MET study, the sample size has decreased to 1306. The schools which provided school context data were used for the multi-group analysis. These two school characteristics, proportion minority and students in poverty were treated as important context variables for high-need schools.

Table 4.8 *Descriptive Statistics and Correlations for School Context Items*

	School Context	Mean	S.D.	1	2	3
1	% Minority	.764	.236	1		
2	% Free and Reduced Meals Program	.680	.246	.368**	1	

N=1306

Note: **, p<.01; *, p<.05

Table 4.8 shows the descriptive statistics of the school context variables. In the sample schools, a large proportion (>50%) of students were identified as minority or poverty, on average. This could be the reason that more disadvantaged schools are more likely to be volunteered to participate in the MET project to receive a bonus or extra funding. (MET had provided extra funding to participating schools).

Exploratory Factor Analysis

Exploratory factor analysis is a method that suggests the number of underlying factors based on empirical analysis. A Maximum Likelihood method was used to arrive at a more parsimonious conceptual understanding of a set of measured variables by determining the number and nature of common factors needed to account for the pattern of correlations among the measured variables. Moreover, an oblique solution (i.e. direct oblimin) was selected to extract factors for latent variables: student perceptions of teaching, working conditions, and school context, because the factors within each latent variable should theoretically be correlated with each other. Each analysis was run separately. Scree plots and total explained variance (Table 4.9) were used to determine the number of factors. The results of the final exploratory factor analysis of each scale were presented in Table 4.10 and Table 4.11. The results of the analysis showed that one single factor can explain student perception. And working conditions can be explained by five factors. For each scale, more than 60% of total variance was explained by the extracted factors.

Table 4.9 *Results of Exploratory Factor Analyses for Scales*

Scale	Number of items	Factors extracted	Explained variance	Total explained variance
Student perception	10	1	73.80%	73.80%
Working conditions	28	5	36.89%	65.76%
			15.13%	
			4.96%	
			4.90%	
			3.82%	

Table 4.10 *EFA for Student Perception Indicators*

Observed indicators	Student perception
My teacher in this class makes me feel that he/she truly cares about me.	.880
My teacher knows if something is bothering me.	.837
My teacher explains difficult things clearly.	.913
My teacher has several good ways to explain each topic that we cover in class.	.924
My teacher knows when the class understands, and when we do not.	.879
If you don't understand something, my teacher explains it another way.	.897
In this class, my teacher accepts nothing less than our full effort.	.778
My teacher checks to make sure we understand what s/he is teaching us.	.924
My teacher wants me to explain my answers—why I think what I think.	.736
My teacher takes the time to summarize what we learn each day.	.800

For working condition indicators, the EFA results confirmed the conceptual structure of all factors. All the indicators showed clear and strong fit with each factor. The indicators that are conceptually related to each of the working condition constructs were loaded together, with strong loadings (>.5). A factor loading of .40 or above is considered to be meaningful (Floyd & Widaman, 1995). There were no cross-loadings, loading on more than one factor, existed for items; hence all items from the survey were retained.

Table 4.11 *EFA for Working Condition Indicators*

Observed indicators	Factors				
	Managing student conduct	Instructional practice	Teaching workload	Instructional resources	Instruction autonomy
Managing student conduce items					
Students at this school understand expectations for their conduct.	**-.620**	.079	-.058	.065	.036
Students at this school follow rules of conduct.	**-.656**	-.012	.074	.081	-.006
School administrators support teachers' efforts to maintain discipline in the classroom.	**-.823**	-.048	.090	-.028	.061
School administrators consistently enforce rules for student conduct.	**-.917**	-.037	.081	-.035	-.043
Policies and procedures about student conduct are clearly understood by the faculty.	**-.614**	.129	-.061	.011	.051
Instructional practice items					
Teachers use assessment data to inform their instruction.	-.029	**.591**	-.077	.054	.007
Teachers work in professional learning communities to develop and align instructional practices.	-.001	**.827**	.046	-.004	-.049
Provided supports (i.e. instructional coaching, professional learning communities, etc.) translate to improvements in instructional practices by teachers.	-.021	**.691**	.099	.007	.047
Teachers are encouraged to try new things to improve instruction.	-.094	**.424**	.101	-.035	.259
Teaching workload items					
Teachers have sufficient instructional time to meet the needs of all students.	.059	.132	**.580**	.059	-.047
Teachers are allowed to focus on educating students with minimal interruptions.	-.180	.016	**.599**	.025	-.024
Efforts are made to minimize the amount of routine paperwork teachers are required to do.	-.044	-.035	**.589**	.030	.140
Teachers are protected from duties that interfere with their essential role of educating students.	-..080	-.059	**.616**	.054	.063
Instructional resource items					
Teachers have sufficient access to appropriate instructional materials.	-.078	.037	.051	**.539**	.073
Teachers have sufficient access to instructional technology, including	.072	.014	-.032	**.922**	-.041

computers, printers, software and internet access.					
Teachers have access to reliable communication technology, including phones, faxes and email.	-.120	.013	-.040	**.681**	-.011
The reliability and speed of Internet connections in this school are sufficient to support instructional practices.	.045	-.022	.103	**.529**	.055
Classroom autonomy					
Teachers are relied upon to make decisions about instruction issues.	-.010	-.003	-.028	.051	**.901**
The faculty has an effective process for making group decisions to solve problems.	-.141	.129	.017	.094	**.456**
Teachers are trusted to make sound professional decisions about instruction.	-.048	-.012	.005	.048	**.851**
Teachers have autonomy to make decisions about instructional delivery (i.e. pacing, materials and pedagogy).	-.039	.117	.163	-.007	**.401**

Structural Equation Modeling

Overview of the Methodology

Structural equation modeling (SEM) is an especially appropriate method for analyzing non-experimental data. In addition to parameter estimates, the programs such as *LISREL* provide fit indices to assess how well the model fits the data. Such fit indices make it possible to evaluate the adequacy of the theoretical model in explaining the data (Bollen, 1989; Schumacker & Lomax, 2010). In this study, the latent variable structural equation models were estimated in several steps separately for ELA and Math teachers, by using LISREL 8.8 computer program (Jöreskog & Sörbom, 2006). First, the measurement of each scale of the model (constructs and their indicators) was specified and estimated; second, the full measurement model was tested and modified; third, structural relationships in the model were specified and estimated. For the purpose of cross-validation, the estimated models were examined and tested for its robustness and accuracy across a second independent sample. The full sample was randomly divided into two groups. The first calibration sample was used to build the model, and the second validation sample was used to validate the model. An invariance testing strategy was used to test for replicability of the measurement models and full structural models across calibration and validation samples. In addition, a multi-group analysis was conducted to examine whether high-need school context plays a role in the relationships.

Measurement Models

Following the results of the exploratory factor analysis, we conducted confirmatory factor analyses to test the measurement models and assess the construct validity of the various latent constructs. The model was constructed by using the calibration sample in three steps. First, the model for each construct was tested separately. Second, the full model was tested for all constructs simultaneously. Third, we evaluated the full measurement model following Fornell and Larcker's (1981) criteria. Then the validation sample was used to validating the estimated measurement model, an invariance testing strategy was used to test for replicability of the measurement models across calibration and validation samples. All models were estimated using the maximum likelihood method.

Measurement of the Student Perceptions of Teaching

Table 4.12 *Goodness-of-Fit Summary Table for Measurement Models of Student Perception Items*

	χ^2	*df*	$\Delta\chi^2$	Δ *df*	CFI	GFI	RMSEA	SRMR
Initial Model	787.89	35			.98	.93	.10	.026
TD (2,1)	550.46	34	237.43	1	.99	.95	.09	.022
TD (9,8)	447.80	33	71.03	1	.99	.96	.078	.020
TD (6,4)	377.30	32	70.50	1	.99	.96	.072	.017
TD (7,5)	311.58	31	65.72	1	.99	.97	.067	.015
TD (4,3)	252.60	30	58.98	1	.99	.98	.060	.014
TD (7,4)	203.06	29	49.54	1	1.00	.98	.053	.012

Student perception about teaching was measured by 10 indicators and all these loadings were significant. To get the better fitting model (CFI=1.00, GFI=.98, AGFI=.93, RMSEA=.05), six pairs of covariance of indicators were set free. Correlating the errors made sense because these items were conceptually related and their errors had some common variances. Chi-square difference test showed the model significantly improved at each step (Table 4.12). Bollen (1989) suggested 3 to 4 indicators are the best number of indicators for each latent construct, so the new composites of correlated indicators were created to simplify and optimize the model.

Measurement of Teacher Value-added Score and Teacher Background Quality

The estimated teacher value-added score was based on a single outcome measured by the state mathematics and English Language Arts (ELA) tests, separately. And teacher background quality was measured by a composite score of teachers' education level and teaching experience. Since teacher value-added score and teacher background quality were the two observed variables in the model; the random errors are fixed in the model.

Measurement of Instructional Practice Support

Table 4.13 *Goodness-of-Fit Summary Table for Measurement Models of Instructional Support Items*

	χ^2	*df*	$\Delta\chi^2$	Δ *df*	CFI	GFI	RMSEA	SRMR
Initial Model	54.09	2			.98	.99	.11	.026
TD (6,5)	.89	1	1	1	1.00	1.00	.00	.004

Instructional practice support was measured by 4 indicators and all these loadings were significant. To get the better fitting model, one pair of covariance of indicators was set free. Chi-square difference test showed the model significantly improved (Table 4.13). A new composite of the average of correlated items was created for future analyses.

Measurement of Teaching Workload

There were four measures of the teaching workload. These 4 indicators showed a well-fitting model (CFI=.99, GFI=.99, AGFI=.97, RMSEA=.07), and all four loadings were significant. Chi-square difference test showed the model was not significantly improved by freeing one pair of covariance of indicators. No item parceling is needed to further improve the model fit.

Measurement of Instruction Resources

Instruction resources was measured by 4 indicators and all these loadings were significant. These 4 indicators showed a well-fitting model (CFI=.99, GFI=.99, AGFI=.99, RMSEA=.051), and all four loadings were significant. Similarly as the indicators of teaching workload, the correlation between indicators was not significantly strong. Hence, no item parceling is needed to further improve the model fitness.

Measurement of Classroom Autonomy

There were four measures of the classroom autonomy. These 4 indicators showed a well-fitting model (CFI=1.00, GFI=1.00, AGFI=.98, RMSEA=.056), and all four loadings were significant. Chi-square difference test showed the model was not significantly improved by freeing one pair of covariance of indicators. No item parceling is needed to further improve the model fit.

Measurement of Support for Managing Student Conduct

Table 4.14 *Goodness-of-Fit Summary Table for Measurement Models of Managing Student Conduct Items*

	χ^2	*Df*	$\Delta\chi^2$	Δ *df*	CFI	GFI	RMSEA	SRMR
Initial Model	261.90	5			.96	.95	.16	.045
TD (4,3)	108.04	4	153.86	1	.99	.98	.11	.027

Support for managing student conduct was measured by 5 indicators and all these loadings were significant. Two indicators were highly correlated. The nested models were tested using chi-square difference tests (Table 4.14) to make the model fit better (GFI=.99, CFI=.98, AGFI=.93, and RMSEA=.11). New composites of the average of correlated items were created for future analyses.

Full Measurement Model

After creating new composites by taking an average of highly correlated indicators, the full hypothesized measurement model was tested, separately for ELA and Math teachers (Figure 4.2). Overall correlation matrices of all scales for ELA and Math teachers are listed in Table 4.15. In the measurement models, the first indicator was set to a value of 1.0 for each latent variable, and thus *t* scores were not computed. The fit indices for both measurement models were high, indicating a well-fitting model in which data fit well to the hypothesized model, $\chi^2(249, N=698) = 698.04$, $p<.05$ for ELA model; and $\chi^2(249, N=644) = 761.47$, $p<.05$ for Math model. The goodness-of-fit indices (GFI) were .93 and .92, and the adjusted goodness-of-fit (AGFI) indices were .91 and .90; the comparative fit indices (CFI) were .97 and .97; the root-mean-square errors of approximation (RMSEA) were .051 and .057; the standardized root mean square errors of approximation (SRMR) were .045 and .045, for ELA and Math model respectively. Overall, these fit indices indicated a theoretically sound model that explained the data well.

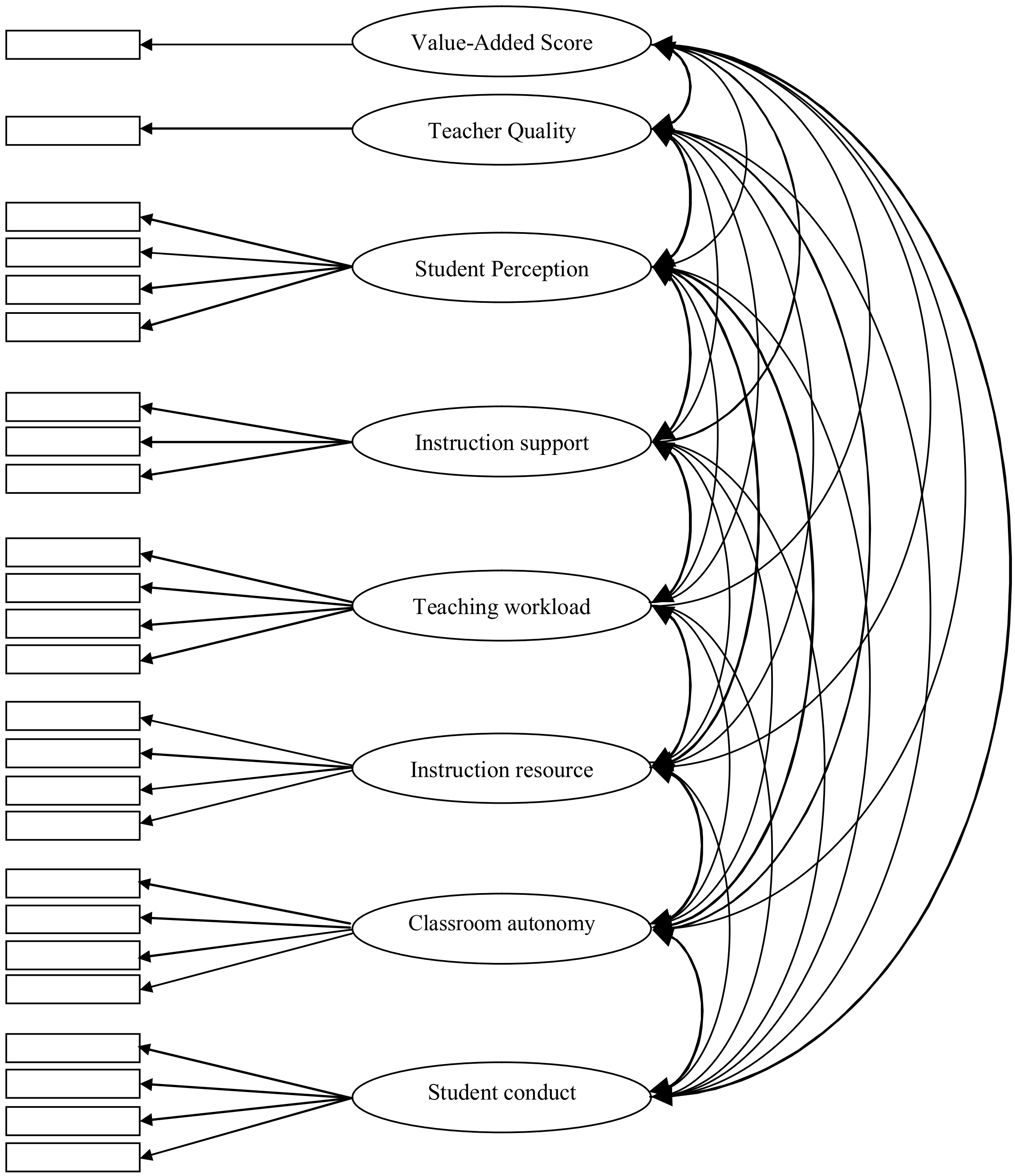

Figure 4.2 Full Measurement Model

Table 4.15 *Overall Item Correlation Matrix (ELA)*

	TVA	TQ	SP1	SP2	SP3	SP4	MSC1	MSC2	MSC3	MSC4	IP1	IP2	IP3
TVA	1												
TQ	.011	1											
SP1	.092	.091	1										
SP2	.132	.082	.844	1									
SP3	.187	.080	.660	.769	1								
SP4	.098	.029	.678	.739	.592	1							
MSC1	.071	.150	.068	.057	.079	-.032	1						
MSC2	.071	.125	.138	.117	.171	-.047	.573	1					
MSC3	.068	.131	.094	.053	.103	-.018	.601	.661	1				
MSC4	.073	.207	.079	.053	.074	.001	.574	.493	.614	1			
IP1	.056	.107	.106	.097	.128	.069	.244	.182	.236	.292	1		
IP2	.075	.125	.001	.002	.043	-.021	.343	.317	.419	.391	.531	1	
IP3	.070	.147	.063	.066	.079	.019	.274	.327	.402	.350	.439	.689	1
TW1	-.024	.068	-.088	-.067	-.067	-.068	.195	.245	.286	.211	.163	.299	.301
TW2	.040	.134	.015	.014	.031	-.034	.329	.433	.472	.329	.154	.321	.314
TW3	-.003	.005	-.032	-.025	-.038	-.026	.243	.305	.406	.263	.069	.279	.260
TW4	.030	.062	.036	.017	-.002	-.016	.268	.332	.417	.271	.139	.272	.282
IR1	.041	.133	.066	.042	.080	-.017	.347	.393	.411	.348	.234	.353	.349
IR2	.049	.104	.074	.092	.071	.052	.250	.312	.319	.269	.224	.277	.292
IR3	.010	.053	.032	.042	.050	.016	.269	.284	.353	.278	.257	.329	.291
IR4	-.003	-.013	-.009	-.037	-.067	-.020	.154	.214	.231	.172	.097	.206	.214
CA1	.027	.032	.061	.022	.026	.019	.376	.396	.479	.399	.266	.488	.439
CA2	.030	.117	.089	.043	.047	.015	.416	.476	.582	.456	.265	.499	.478
CA3	.035	.053	.070	.038	.027	.020	.366	.404	.506	.398	.242	.475	.426
CA4	.014	.062	.012	.012	.025	.018	.232	.278	.360	.260	.242	.438	.370

(Continued)

	TW1	TW2	TW3	TW4	IR1	IR2	IR3	IR4	CA1	CA2	CA3	CA4
TW1	1											
TW2	.463	1										
TW3	.389	.474	1									
TW4	.361	.493	.514	1								
IR1	.286	.362	.319	.332	1							
IR2	.260	.277	.258	.307	.549	1						
IR3	.247	.276	.254	.271	.484	.607	1					
IR4	.251	.230	.244	.243	.311	.482	.442	1				
CA1	.290	.365	.407	.382	.430	.329	.332	.273	1			
CA2	.301	.430	.368	.349	.412	.369	.345	.279	.686	1		
CA3	.269	.389	.405	.409	.399	.333	.331	.272	.837	.657	1	
CA4	.300	.322	.405	.335	.294	.230	.225	.206	.519	.423	.537	1

Overall Item Correlation Matrix (Math)

	TVA	TQ	SP1	SP2	SP3	SP4	MSC1	MSC2	MSC3	MSC4	IP1	IP2	IP3
TVA	1												
TQ	.037	1											
SP1	.110	.061	1										
SP2	.139	.072	.860	1									
SP3	.147	.072	.692	.764	1								
SP4	.061	-.010	.677	.740	.607	1							
MSC1	.088	.152	.079	.059	.097	-.026	1						
MSC2	.131	.082	.168	.135	.165	-.033	.617	1					
MSC3	.101	.055	.072	.052	.099	-.005	.557	.648	1				
MSC4	.036	.169	.053	.039	.086	.012	.533	.444	.571	1			
IP1	.067	.089	.109	.106	.134	.097	.258	.209	.255	.318	1		
IP2	.086	.078	.004	.007	.062	-.021	.341	.321	.421	.393	.544	1	
IP3	.061	.068	.067	.073	.073	.046	.307	.325	.379	.343	.448	.662	1
TW1	.046	-.018	-.085	-.038	-.019	-.036	.194	.231	.286	.173	.171	.293	.309
TW2	.078	.043	-.003	.013	.067	-.052	.353	.402	.441	.324	.156	.321	.278
TW3	.055	-.081	-.049	-.036	-.024	-.022	.232	.281	.392	.225	.080	.314	.259
TW4	.068	.019	.041	.043	.041	.004	.248	.314	.406	.268	.178	.320	.265
IR1	.017	.108	-.032	-.042	-.026	-.071	.278	.300	.317	.266	.246	.329	.301
IR2	.062	.062	.028	.030	.011	-.002	.198	.257	.241	.209	.265	.302	.281
IR3	.044	-.002	.062	.036	-.012	.016	.202	.258	.339	.221	.245	.318	.260
IR4	.056	-.114	-.027	-.040	-.060	-.047	.147	.207	.239	.147	.115	.261	.244
CA1	.009	.029	.033	.029	.036	.013	.360	.371	.454	.349	.258	.492	.385
CA2	.037	.044	.065	.030	.037	.021	.379	.425	.569	.420	.267	.515	.436
CA3	-.011	.003	.012	.015	.035	.005	.355	.392	.491	.362	.264	.514	.403
CA4	.034	.011	.098	.059	.076	.033	.182	.262	.314	.212	.230	.450	.354

(Continued)

	TW1	TW2	TW3	TW4	IR1	IR2	IR3	IR4	CA1	CA2	CA3	CA4
TW1	1											
TW2	.436	1										
TW3	.398	.430	1									
TW4	.420	.522	.492	1								
IR1	.272	.270	.263	.273	1							
IR2	.264	.229	.229	.291	.543	1						
IR3	.223	.226	.237	.263	.418	.593	1					
IR4	.237	.237	.245	.222	.287	.491	.433	1				
CA1	.305	.361	.406	.365	.338	.276	.293	.252	1			
CA2	.290	.446	.388	.367	.341	.306	.343	.274	.661	1		
CA3	.310	.376	.435	.385	.358	.278	.300	.268	.822	.632	1	
CA4	.309	.335	.445	.328	.220	.243	.249	.231	.456	.396	.503	1

Table 4.16 *Standardized Loading, Reliability, and Validity of the Final Measurement Model*

Label	Construct and indicators	**ELA model**				**Math model**			
		Standardized loading	t	Reliability	Variance extracted estimate	Standardized loading	t	Reliability	Variance extracted estimate
TVA	Teacher value-added score	.74	--	.55	.55	.88	--	.78	.78
TQ	Teacher background quality	.98	--	.97	.97	.98	--	.97	.97
	Student perception about teaching			.90	.71			.90	.75
SP1	My teacher in this class makes me feel that s/he really cares about me. +My teacher seems to know if something is bothering me	.85	--	.72		.89	--	.79	
SP2	My teacher explains difficult	.97	35.27	.94		.97	39.57	.94	

	things clearly. +My teacher has several good ways to explain each topic that we cover in this class. +My teacher knows when the class understands, and when we do not. +If you don’t understand something, my teacher explains it another way. +My teacher checks to make sure we understand what s/he is teaching us.								
SP3	In this class, my teacher accepts nothing less than our full effort. +My teacher wants me to explain my answers—why I think what I think.	.80	26.71	.64		.80	27.23	.64	
SP4	My teacher takes the time to summarize what we learn each day.	.74	23.27	.55		.78	26.14	.61	
	Managing student conduct			.85	.60			.87	.54
MSC1	Students at this school understand expectations for their conduct.	.73	--	.53		.73	--	.53	
MSC2	Students at this school follow rules of conduct.	.81	20.30	.66		.74	17.03	.55	
MSC3	School administrators support teachers' efforts to maintain discipline in the classroom. +School administrators	.87	21.59	.76		.78	17.82	.61	

	consistently enforce rules for student conduct.								
MSC4	Policies and procedures about student conduct are clearly understood by the faculty.	.68	17.06	.46		.67	15.61	.45	
	Instructional practice			.79	.58			.80	.60
IPS1	Teachers use assessment data to inform their instruction.	.62	--	.38		.62	--	.38	
IPS2	Teachers work in professional learning communities to develop and align instructional practices.	.88	15.90	.77		.90	15.84	.81	
	+Teachers are encouraged to try new things to improve instruction.								
IPS3	Provided supports (i.e. instructional coaching, professional learning communities, etc.) translate to improvements in instructional practices by teachers.	.77	15.59	.59		.78	15.42	.61	
	Teaching workload			.77	.44			.78	.46
TW1	Teachers have sufficient instructional time to meet the needs of all students.	.53	--	.28		.62	--	.38	
TW2	Teachers are allowed to focus on educating students with minimal interruptions.	.76	12.64	.58		.70	13.19	.49	
TW3	Efforts are made to minimize the amount of routine paperwork teachers are required to do.	.67	11.94	.45		.65	12.64	.41	

TW4	Teachers are protected from duties that interfere with their essential role of educating students.	.67	11.94	.45		.75	13.67	.56	
	Instruction resources			.77	.49			.76	.50
RF1	Teachers have sufficient access to appropriate instructional materials.	.69	--	.48		.71	--	.50	
RF2	Teachers have sufficient access to instructional technology, including computers, printers, software and internet access.	.78	15.58	.61		.79	15.50	.63	
RF3	Teachers have access to reliable communication technology, including phones, faxes and email.	.72	15.33	.52		.70	13.68	.49	
RF4	The reliability and speed of Internet connections in this school are sufficient to support instructional practices.	.58	12.22	.34		.60	12.72	.36	
	Classroom autonomy			.83	.52			.84	.56
CA1	Teachers are relied upon to make decisions about instruction issues.	.73	--	.53		.83	--	.69	
CA2	The faculty has an effective process for making group decisions to solve problems.	.76	14.82	.58		.78	13.86	.61	
CA3	Teachers are trusted to make sound professional decisions about instruction.	.79	13.83	.63		.80	13.14	.64	

CA4	Teachers have autonomy to make decisions about instructional delivery (i.e. pacing, materials and pedagogy).	.57	13.42	.32		.55	11.26	.30

The standardized item loadings, item reliability, construct reliability and variance extracted estimate were evaluated according to Fornell and Larcker's (1981) criteria (Table 4.16). A variance extracted of greater than 0.50 indicates that the validity of both the construct and the individual variable is high. The variance extracted the estimate of student perception was very high (over .70) for both ELA and Math models. For two working condition latent constructs, Teaching workload and Instruction resource, the estimates were slightly lower than .50. But it is important to note that this variance extracted estimate test is very conservative. Given the significant factor loadings and high reliabilities, the constructs were retained in the final measurement models.

Validating the Measurement Model

The measurement models were already established as baseline models for the cross-validation by using the calibration sample. Then the invariance testing strategy was used to test for replicability of the measurement models across calibration and validation groups. The invariance testing was performed by comparing the relevant parameters across groups using chi-square. First, an unconstrained model that factor loadings were allowed to vary freely but with same factor structure across groups was estimated. Then, a joint constrained model was estimated where the factor loadings across groups are constrained to be equal to each other. Last, the fit of the constrained model with that of an unconstrained one were compared using the difference in chi-square statistic with a significance level of .05. If the chi-square difference statistic does not reveal a significant difference between the unconstrained and constrained models, then we can conclude that the constrained model applies across calibration and validation groups.

Table 4.17 *Results of Test of Invariance of Measurement Model for Cross-validation*

Measurement Model	χ^2	*df*	*p* value	χ^2/df	RMSEA	CFI
ELA teacher model						
Model A (unconstrained)	1490.49	557	<.01		.049	.96
Model B (constrained)	1513.94	574	<.01		.048	.96
$\chi^2_{difference}$ (Model B-Model A)	23.45	17	>.05			
Math teacher model						
Model A (unconstrained)	1597.53	557	<.01		.054	.95
Model B (constrained)	1619.49	574	<.01		.053	.95
$\chi^2_{difference}$ (Model B-Model A)	21.96	17	>.05			

Results of tests of invariance of the measurement models indicated that factor loadings did not differ between calibration and validation samples for both ELA and Math teachers. The chi-square differences between the unconstrained and constrained models were 23.45 (*df*=17) in ELA teacher model, and 21.96 (*df*=17) in Math teacher model (See Table 4.17). The associated *p* values were both greater than .5, indicating no difference between the unconstrained and constrained models. Other fit indices for the constrained model showed a better fit than those of unconstrained model. No significant differences in factor loadings between calibration and validation samples suggested that the measurement model validation was successful.

Structural Models

Relationships among Two Measures of Teacher Effectiveness and Teacher Background Quality

The teacher effectiveness model included paths from teacher background quality to teacher value-added score and students' perception of teaching; and paths from student perception to teacher value-added score. ELA and Math teachers were tested separately in two models. Both models had good fit, χ^2=28.26, *df* =8, *p*<.01 for ELA model; and χ^2=19.05, *df* =8, *p*<.01 for Math model; RMSEA=.060, and .045; CFI=.99, and 1.00; GFI=.99, and 99; SRMR=.022, and .021; AGFI=.96, and .98, for ELA and Math model respectively. In the teacher effectiveness model, a very small proportion of variance (R^2 = .03, the same in ELA and Math model) of teacher value-added scores was explained by student perception about teaching and teacher background quality.

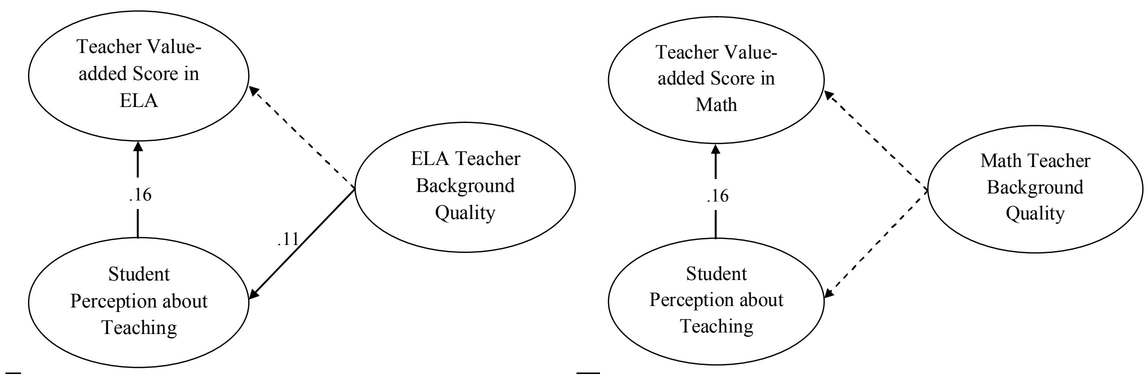

Figure 4.3 Structural Models of Teacher Background Quality, Teacher Value-added Scores, and Student Perception, for ELA and Math Teachers

Both ELA and Math models had similar results in general but there were some differences. Figure 4.3 shows the ELA and Math structural models, the dash lines indicated the effects are not statistically significant. The paths from student perception to teacher value-added scores were significant for ELA and Math teachers. But the path from teacher background quality to student perception was only significant in English language arts. And the path from teacher background quality to teacher value-added score was not significant. These two measures of teacher effectiveness and teacher background quality were positively correlated with each other.

Table 4.18 *Standardized Direct, Indirect, and Total Effects of Teacher Background Qualities on Teacher Value-added Score and Student Perception*

	ELA model	**Math model**

	Student perception	VA ELA score	Student perception	VA Math score
Student perception				
Direct	--	$.16^{*}$	--	$.16^{*}$
Indirect				
Total		$.16^{*}$		$.16^{*}$
Teacher background quality				
Direct	$.11^{*}$	.00	.05	.02
Indirect		$.02^{*}$		.01
Total	$.11^{*}$	.02	.05	.03

Note: *, p<.05

The direct, indirect and total effects in the model were shown in Table 4.18. The effects of teacher background quality were positive and direct on student perception (β=.09 in ELA model; and =.07 in Math model), the effect was significant only for ELA teachers. Although the direct effect of teacher background quality to teacher value-added score was not significant, there was a significant indirect effect only in English language arts. Student perception had a significant effect on teacher value-added scores (β=.16 in ELA and Math) in both English language arts and mathematics. This implied that teacher value-added score was significant positively correlated to their students' perception and their background quality, although the correlations were not strong.

Effects of Working Conditions on Teacher Value-added Scores

The effects of all working conditions on teachers' value-added scores were tested simultaneously. ELA and Math teachers are tested separately in two models. Both models had good fit, χ^2=484.22, df=175, p<.01 for ELA model; and χ^2=609.23, df=175, p<.01 for Math model; RMSEA=.050, and .052; CFI=.97, and .95; GFI=.95, and 93; SRMR=.043, and .049; AGFI=.93, and .90, for ELA and Math model respectively.

Table 4.19 *Standardized Effects of Working Conditions on Teacher Value-added Score*

	Teacher value-added score in ELA	Teacher value-added score in math
Teacher background quality	-.01	.03
Managing student conduct	.08	.05
Instructional support	.15*	.13*
Teaching workload	-.07	.01
Instruction resources	.01	-.07
Classroom autonomy	.01	.02

Note: *, $p<.05$

Both ELA and Math models have similar results (see Table 4.19). Across all working conditions, only instructional practice support had a significant effect on teacher value-added score for both ELA (β=.15) and Math (β=.13) teachers. Although other working conditions, including managing student conduct and classroom autonomy had positive effects on teacher value-added score, the effects were weak and not significant, in either English language arts or mathematics. In addition, the intercorrelations among working conditions were found significant in a range from -.05 to .64. Particularly, support for managing student conduct, instructional support, and teaching workload were strongly correlated to each other. However, these correlations among working conditions did not lead to a severe multicollinearity problem in the models. The multicollinearity was examined according to Grewal, Cote, & Baumgartner's (2004) criteria: no standardized coefficients were found greater than 1 or less than -1; and no negative estimates of variance were found.

Effects of Each Working Condition on Teacher Value-added Scores and Student Perception

To better understand the effects of each working condition, a series of structural models were estimated by regressing teacher value-added score and student perception on each working condition separately. In this step, the effects of each working condition were examined in separate models. Each model included paths from a single working condition to teacher value-added score and to the students' perception about teaching, and paths from teacher background quality to teacher value-added score and the student perception. Table 4.20 shows the fit indices of each model. The fit indices for each model were high, indicating a well-fitting model in which data fit well to the model. The goodness-of-fit indices (GFIs) of all models were high (≥.97), and the adjusted goodness-of-fit (AGFIs) indices were good (≥.94). The comparative fit indices (CFIs) of all models were high as well (≥.97). The standardized root mean square errors of approximation (RMSEAs) and standardized root mean square residuals (SRMRs) of all models were less than .069 and .050, respectively. Similarly, the interactions among teacher background, student perception, and teacher value-added score were very consistent with the results of prior structural models which only include teacher effectiveness and background qualities.

Table 4.20 *Separate Models: Model Fit indices*

	χ^2	CFI	GFI	AGFI	RMSEA	SRMR
ELA Models						
Managing student conduct	211.02	.97	.97	.94	.069	.037
Instructional support	84.53	.98	.97	.95	.060	.050
Teaching workload	88.63	.98	.98	.96	.050	.038
Instruction resources	101.85	.98	.97	.95	.056	.041
Classroom autonomy	65.13	.99	.98	.97	.039	.026
Math Models						
Managing student conduct	153.79	.98	.97	.95	.060	.036
Instructional support	64.76	.99	.98	.96	.049	.040
Teaching workload	94.43	.98	.97	.95	.053	.036
Instruction resources	93.23	.98	.97	.96	.052	.038
Classroom autonomy	60.40	.99	.98	.97	.036	.030

In each ELA model, the paths from teacher background quality composite to student perception were positive and significant. The paths from student perception to teacher value-added scores were significant. Across all ELA models, only support for managing student conduct and instructional practice support had significant effects on the value-added score; and only managing student conduct and classroom autonomy showed significant effects on student perception. Figure 4.4 illustrates the model of managing student conduct in English language arts as an exemplar. Controlling for teacher background quality, the path from managing student conduct to both student perception and teacher value-added score were significant.

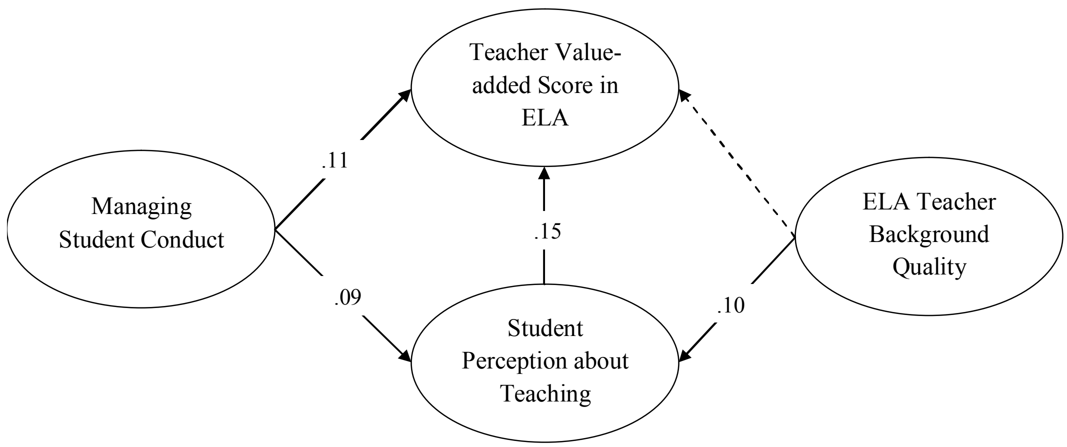

Figure 4.4 Structural Model of Managing Student Conduct on Teacher Value-added Scores and Student Perception, for ELA Teachers

Similar to the results of ELA teacher models, the paths from student perception to teacher value-added scores were significant in each Math teacher model. But the paths from teacher background quality to student perception were not all significant; this effect of teacher background quality to student perception was not significant in managing student conduct and instruction resources models. Across all Math models (the separate models), only instruction practice support and teaching workload had significant effects on teachers' value-added scores in Math, all working conditions but teaching workload and instruction resource showed significant effects on student perception about teaching.

Table 4.21 *Separate Models: Standardized Direct, Indirect, and Total Effects of Each Working Conditions on Teacher Value-added Scores and Student Perception*

	ELA model		Math model	
	Student perception	VA ELA score	Student perception	VA Math score
Direct effects				
Managing student conduct	.11*	.09*	.13*	.06
Instructional support	.04	.15*	.09*	.11*
Teaching workload	-.03	.06	-.02	.07*
Instruction resources	.03	.06	.00	.02
Classroom autonomy	.11*	-.03	.17*	-.09
Indirect effects				
Managing student conduct		.02*		.02*
Instructional support		.01		.01
Teaching workload		-.01		.00
Instruction resources		.01		.00
Classroom autonomy		.02*		.03*
Total effects				
Managing student conduct	.11*	.11*	.13*	.08*
Instructional support	.04	.16*	.09*	.12*
Teaching workload	-.03	.05	-.02	.07*
Instruction resources	.03	.07	.00	.02
Classroom autonomy	.11*	-.01	.17*	-.06

Note: *, p<.05

Table 4.21 provided the standardized direct, indirect and total effects of each working conditions on teacher value-added scores and student perception from the separate models for English language arts and Mathematics. In English language arts, managing student conduct and instructional practice support were directly and positively related to teacher value-added score; managing student conduct and classroom autonomy were significantly and positively related to student perception. In Mathematics, instructional practice support and teaching workload showed significant and direct effects on teacher value-added score; and three of five working conditions had significant and direct effects on student perception. In addition, managing student conduct and classroom autonomy showed significant indirect effects on teacher value-added score in both English language arts and mathematics. Of the five working condition constructs, instructional practice support stood out as a particularly strong predictor to teacher value-added scores of both English language arts and mathematics; support for managing student conduct and classroom autonomy showed strong effects on student perception than other working conditions. Schools that support teachers' development through effective support for managing student conduct in classroom, and provide instructional practice related professional support and resources are more likely to have teachers with higher value-added scores, though it should be noted that the effects are quite small.

Full Model: Effects of All Working Conditions on Teacher Value-added Score and Student Perception

In this step, the full structural model included paths from all working conditions simultaneously to teacher value-added score and to student perception about teaching, and paths from teacher background quality to teacher value-added score and student perception. Both models of the effects of working conditions were fitted well. The ELA model had a good fit, χ^2=732.11, *df* =254, *p*<.01, RMSEA=.052, CFI=.96, GFI=.92, AGFI=.90, and SRMR=.051. The Math model also had a good fit, χ^2=737.32, *df* =254, *p*<.01, RMSEA=.057, CFI=.96, GFI=.91, AGFI=.90, and SRMR=.047. After controlling for teacher background quality, teachers' working conditions explained about 5 percent of the variation in teachers' value-added score and 6 percent of the variation in student perception, the same in English language arts and mathematics.

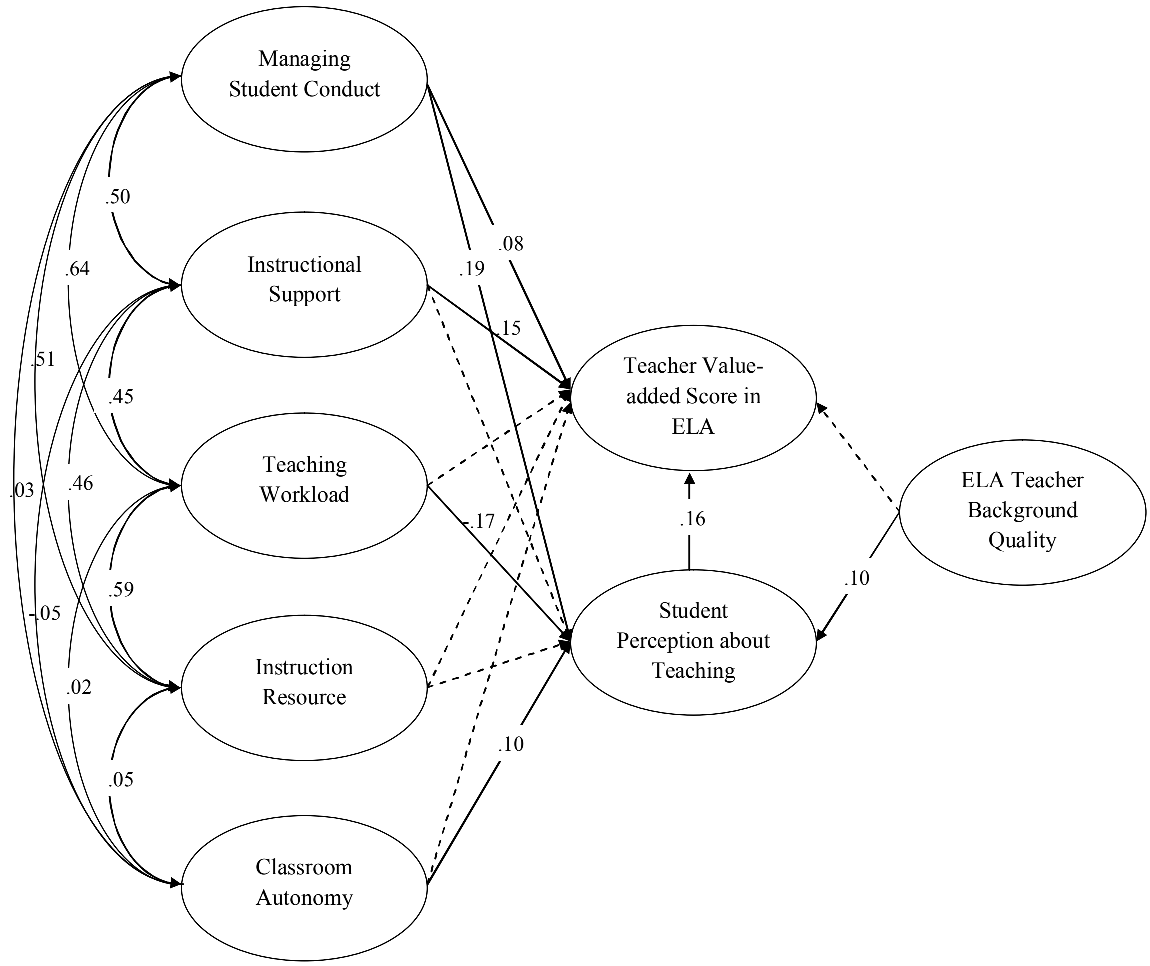

Figure 4.5 Full Structural Model of ELA Teachers

Figure 4.5 illustrated the full structural model of ELA teachers. The paths from managing student (β=.08) and instruction support (β=.15) to teacher value-added score were significant. The paths from the rest working condition constructs to teacher value-added score were not significant. Besides, across all working conditions, support for managing student conduct (β=.19) and classroom autonomy (β=.10) showed significant effects on student perception, and the path from teaching workload (β=.-15) to student perception was significant but negative. Similar to the previous results for ELA teachers, the path from teacher background quality (β=.10) to student perception and the path from student perception (β=.16) to teacher value-added scores did not change after all working conditions were added in the model. All working condition constructs were positively correlated with each other.

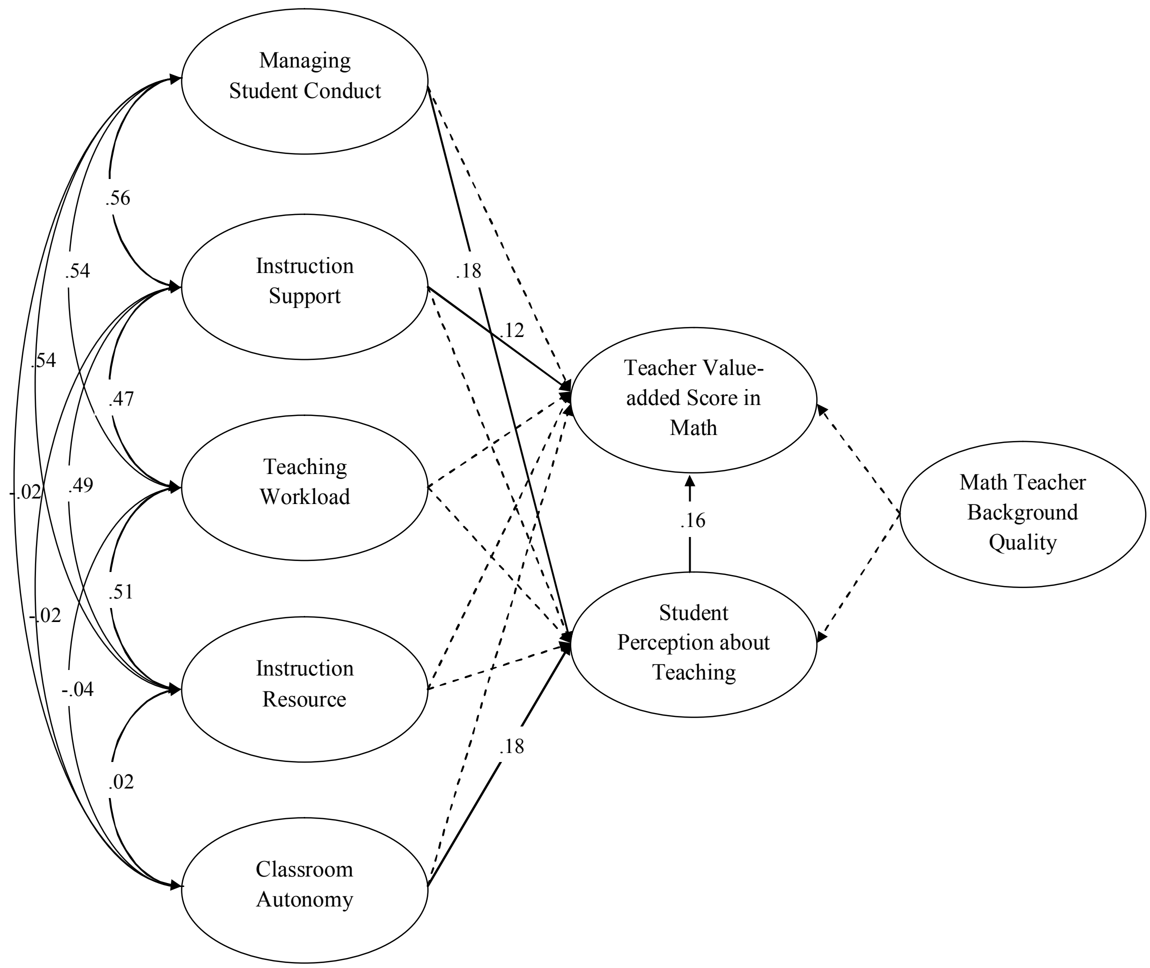

Figure 4.6 Full Structural Model of Math Teachers

Figure 4.6 illustrated the full structural model of Math teachers. Mostly, the paths were similar to the ELA model, but the significant paths to teacher value-added score were different. The only significant path to teacher value-added score was from instruction support (β=.12). Similar to the results in ELA teacher model, managing student conduct (β=.18) and classroom autonomy (β=.18) showed significant and positive effects on student perception. Similar to the results of previous Math teacher model, the path from student perception (β=.16) to teacher value-added scores did not change after all working conditions were added in the model. But the paths from teacher background quality to student perception were not significant. As expected, managing student conduct, instruction support, teaching workload, and instruction resources were strongly correlated with each other.

Table 4.22 *Full Models: Standardized Direct, Indirect, Total Effects of Working Conditions on Teacher Value-added Scores and Student Perception*

	ELA model		**Math model**	
	Student perception	Value-added score	Student perception	Value-added score
Student perception				
Direct		.16*		.16*
Indirect				
Total		.16*		.16*
Teacher quality				
Direct	.10*	-.03	.04	.02
Indirect		.02		.01
Total	.10*	-.01	.04	.03
Managing student conduct				
Direct	.19*	.08*	.18*	.01
Indirect		.03*		.03*
Total	.19*	.11*	.18*	.04
Instructional practice				
Direct	.02	.15*	.08	.12*
Indirect		.00		.01
Total	.02	.15*	.08	.13*
Teaching workload				
Direct	-.17*	-.04	-.08	-.02
Indirect		-.03		-.01
Total	-.17*	-.07	-.08	.01
Instruction resources				
Direct	.02	.01	-.10	-.05
Indirect		.00		-.02
Total	.02	.01	-.10	-.07
Classroom autonomy				
Direct	.10*	.00	.18*	-.07
Indirect		.02		.03*
Total	.10*	.02	.18*	-.05

Note. *, $p<.05$

The direct, indirect and total effects in the model are shown in Table 4.22. Teacher background quality effects were significant and positively correlated to student perception (β=.10) in English language arts, but not significant in mathematics. In both ELA and Math models, teacher background quality was weakly correlated to teacher value-added score; but student perception had a significant effect on teacher value-added score (β=.16 in ELA and Math). Management of student conduct had a significant direct effect on teacher value-added score (β=.08) only for ELA teachers, and it had a significant indirect effect on teacher value-added score for both teachers. Moreover, managing student conduct had an influential direct effect on student perception (β=.19 in ELA; .18 in Math) in both English language arts and mathematics. Instructional practice had a very weak effect on student perception, but it had significant effects on teacher value-added score (β=.12 in ELA; .11 in Math), for both ELA and Math teachers. Teaching workload had a negative effect on student perception (β=-.17 in ELA; -.08 in Math), the effect was significant only for ELA teachers. Instruction resource had weak and non-significant effects on either student perception or teacher value-added score. Classroom autonomy was weakly correlated to teacher value-added score in both models; but it showed significant and positive effects on student perception (β=.10 in ELA; .18 in Math) in both English language arts and mathematics. It also had a strong indirect effect on teacher value-added score, but the effect was significant only for math teachers. Thus, classroom autonomy and managing student conduct had strong indirect effects on teacher value-added score mediated through the students' perceived teaching quality in English language arts and mathematics. Besides, the Pearson correlation coefficients of these working condition constructs were positive and strong. This implies that teachers' perceptions of all working conditions are very consistent.

In summary, among these five types of working conditions, support for instructional practices primarily directly affected teacher value-added score. Managing student conduct and classroom autonomy directly affected student perception about teaching, and they also showed indirect effects on teacher value-added score. It is possible that managing student conduct and classroom autonomy affected teacher value-added score through student perception. These effects on teacher value-added score or student perception were slightly different between ELA and Math teachers. In addition, student perception of teaching was a significant predictor of teacher value-added score.

Validating the Structural Model

The structural models were already established as baseline models for the cross-validation by using calibration sample. Then the invariance testing strategy was used to test for replicability of the structural models across calibration and validation groups. The invariance testing was performed by comparing the relevant parameters across groups using chi-square. First, an unconstrained model that coefficients were allowed to vary freely across groups was estimated. Then, a joint constrained model was estimated where the path coefficient across groups is constrained to be equal to each other. Last, the fit of the constrained model with that of an unconstrained one were compared using the difference in chi-square statistic with a significance level of .05. If the chi-square difference statistic does not reveal a significant difference between the unconstrained and constrained models, then we can conclude that the constrained model applies across calibration and validation groups.

Table 4.23 *Results of Test of Invariance of Structural Model for Cross-validation*

Structural Model	χ^2	*df*	*p* value	χ^2 /*df*	RMSEA	CFI
ELA teacher model						
Model A (unconstrained)	1533.62	548	<.01		.051	.96
Model B (constrained)	1538.61	561	<.01		.050	.96
$\chi^2_{difference}$ (Model B-Model A)	4.99	13	>.05			
Math teacher model						
Model A (unconstrained)	1735.97	548	<.01		.056	.95
Model B (constrained)	1743.93	561	<.01		.055	.95
$\chi^2_{difference}$	7.96	13	>.05			

(Model B-Model A)

Results of tests of invariance of the structural models indicated that path coefficients did not differ between calibration and validation samples for both ELA and Math teachers. The chi-square differences between the unconstrained and constrained models were 4.99 (*df*=13) in ELA teacher model, and 7.96 (*df*=13) in Math teacher model (See Table 4.23). The associated *p* values were both greater than .5, indicating no difference between the unconstrained and constrained models. No significant differences in factor loadings between calibration and validation samples suggested that the structural model validation was successful.

Multi-Group SEM Analyses

Multi-Group SEM analyses were performed to examine whether structural relationships in the full working condition structural model differed across high-need school context. The full sample included school context information (*N*=1306) was used for the multi-group analysis. As a prerequisite to testing for differences in the strength of the structural relationships, it is customary to first establish a baseline model for each subgroup separately. First, the equivalence of the measurement model was established, and second, the structural models were compared (Byrne, 1998; Joreskog & Sorbom, 1993; Kelloway, 1998).

Given that there was no significant difference between the unconstrained and constrained or partially constrained measurement models (i.e., invariant factor loadings), the structural paths of interest among the latent variables were compared by examining chi-square differences and other fit indices (e.g., χ^2 */df* and RMSEA) between the fully unconstrained model and the model with a constrained path of interest. In comparing this partially constrained model with the unconstrained one, the differences in chi-squares were examined to evaluate whether the fit of the constrained model is significantly worse than that of the unconstrained. The subgroups for multi-group analysis were drawn from three high-need school context discussed in the previous section.

Percent of Minority Students

Some studies on teacher effectiveness have pointed out that high-need schools with higher proportions of minority students are hard to staff as working environment is less desirable and thus, these schools tend to have less effective teachers (Aaronson, Barrow, & Sander, 2007; Johnson, Kraft, & Papay, 2012). To examine the impact of the percent of minority students on the structural relationships, this study created two groups based on 1306 teachers with school context information: for schools in the top quartile of minority students enrollment; and for schools in the bottom quartile of minority students enrollment. Table 4.24 presented the results of statistical tests that compare differences in means or percentages. Comparison of means was based on Tamhane's T2, which does not assume equal variances across groups.

Table 4.24 *Descriptive Statistics for All Variables by Percent of Minority Students*

	Low Minority		High Minority		Sig.
	Mean	(S.D.)	Mean	(S.D.)	
Teacher value-added score in ELA	.026	.247	-.003	.162	
Teacher value-added score in Math	.053	.332	-.010	.170	*
Teacher background quality	2.24	1.051	2.11	1.123	
Student perception	4.05	.402	3.79	.380	***
Working conditions					

Managing student conduct	3.13	.583	2.66	.726	***
Instructional practice	3.20	.524	3.06	.535	**
Teaching workload	2.75	.572	2.45	.648	***
Instruction resources	3.12	.588	2.85	.631	***
Classroom autonomy	3.06	.611	2.76	.650	***

Note. Top and bottom quartiles of schools by the proportion of minority students. The top quartile is 97.84 percent or greater minority; the bottom quartile is 57.42 percent or fewer minority students.
**p*<0.05; ** *p*<0.01; ****p*<0.001.

Teachers in schools with high proportion of minority students had significantly lower value-added score in mathematics, though differences in English language arts were not significant. Teachers in schools serving a high proportion of minority students had worse students' evaluation of their teaching in classroom. The findings with regard to teachers' perception data generally suggested that schools serving fewer minority students have better working conditions. In the schools with fewer minority students, teachers expressed more positive perceptions of support in managing student conduct, instructional practice, teaching workload, instruction resource, and classroom autonomy.

Before testing whether the structural relationships were different, the equivalence of the measurement model was tested. As indicated in Table 4.25, the chi-square difference between the constrained and unconstrained models was 41.21 with 17 degrees of freedom ($p<.01$) for ELA teachers and 33.26 with 17 degrees of freedom ($p<.01$) for Math teachers, indicating there was a difference between the unconstrained and constrained models for both ELA or math teachers. In other words, the unconstrained model fitted the data more closely than the constrained model. These results indicated that factorial invariance did not exist. Next, an attempt was made to find out which factor loadings actually differed. To do that, modification indices were examined to assess which paths should be estimated separately for the different groups.

Table 4.25 *Results of Test of Invariance of Measurement Model for Percent of Minority*

Measurement Model	χ^2	*df*	*p* value	χ^2/df	RMSEA	CFI
ELA teacher model						
Model A (unconstrained)	1252.47	557	<.01	2.25	.063	.96
Model B (constrained)	1293.68	574	<.01	2.25	.063	.96
Model C (partial constrained)	1277.36	573	<.01	2.23	.063	.96
$\chi^2_{difference}$ (Model B-Model A)	41.21	17	<.01			
$\chi^2_{difference}$ (Model C-Model A)	24.89	16	>.05			
Math teacher model						
Model A (unconstrained)	1314.71	557	<.01	2.36	.061	.95
Model B (constrained)	1347.97	574	<.01	2.35	.061	.95
Model C (partial constrained)	1337.21	573	<.01	2.33	.060	.95
$\chi^2_{difference}$ (Model B-Model A)	33.26	17	<.01			
$\chi^2_{difference}$ (Model C-Model A)	22.50	16	>.05			

For ELA teacher model, modification indices suggested that allowing the factor loading for item SP3 (composite of "In this class, my teacher accepts nothing less than our full effort." and "My teacher wants me to explain my answers—why I think what I think.") on student perception to vary would improve the fit of the constrained model. The same modification made for Math teacher model to improve the fit of the constrained model. After free estimation of factor loading for this item, the partially constrained model was estimated again to compare that model to the fully unconstrained model using the chi-square difference statistic. The difference between this partially constrained model and unconstrained model for ELA teacher was 24.89 with 16 degrees of freedom and an associated p value .072 (>.05), indicating that the partially constrained model and the unconstrained model did not differ significantly. The difference between this partially constrained model and unconstrained model for Math teacher was 22.50 with 16 degrees of freedom and an associated p value .128 (>.05), indicating that the partially constrained model and unconstrained model were the same.

After free estimation of factor loading for item SP3, the statistical significance of the differences in path coefficients was tested. As indicated in Table 4.26, the impact of the percentage of minority students had a significant impact on the structural path coefficients among variables for ELA and Math teachers. The chi-square difference between the constrained and unconstrained models was 25.61 with 13 degrees of freedom (p<.05) for ELA teachers; and 23.90 with 12 degrees of freedom (p<.01) for Math teachers. This result indicated there was a difference between the unconstrained and constrained for ELA and Math teachers. Then, I attempted to find out which path coefficient actually differed.

Table 4.26 *Results of Test of Invariance of Structural Model for Percent of Minority Students*

Structural Model	χ^2	*df*	*p* value	χ^2 /*df*	RMSEA	CFI
ELA teacher model						
Model A (unconstrained)	1228.97	547	<.01	2.25	.063	.96
Model B (constrained)	1254.58	560	<.01	2.24	.063	.96
Model C (partial constrained)	1238.36	558	<.01	2.22	.062	.96
$\chi^2_{difference}$ (Model B-Model A)	25.61	13	<.05			
$\chi^2_{difference}$ (Model C-Model A)	9.39	11	>.05			
Math teacher model						
Model A (unconstrained)	1242.59	547	<.01	2.27	.069	.95
Model B (constrained)	1266.49	560	<.01	2.26	.068	.95
Model C (partial constrained)	1253.12	559	<.01	2.24	.068	.95
$\chi^2_{difference}$ (Model B-Model A)	23.90	13	<.05			
$\chi^2_{difference}$ (Model C-Model A)	10.53	12	>.05			

For ELA teacher model, modification indices suggested that allowing the path from managing student conduct to student perception and from student perception to teacher value-added score to vary would improve the fit of the constrained model. For Math teacher model, modification indices suggested that allowing the path from managing student conduct to student perception to vary would improve the fit of the constrained model. After free estimation of the coefficients for these paths, the partially constrained model was estimated again to compare that model to the fully unconstrained model using the chi-square difference statistic. The difference between this partially constrained model and unconstrained model for ELA teacher was 9.39 with 11 degrees of freedom and an associated p value .586 (>.05). The difference between this partially constrained model and unconstrained model for Math teacher was 10.53 with 12 degrees of freedom and an associated p value .569 (>.05). These results indicated that the partially constrained model and the unconstrained model did not differ significantly. It also suggested that percentage of minority students in schools where teachers taught made a significant difference in several associations among the variables of the study.

Table 4.27 *Structural Path Coefficients between Low and High Percent of Minority Students*

Paths		Standardized path coefficients			
		ELA teachers		**Math teachers**	
From	To	Low Minority	High Minority	Low Minority	High Minority
Student perception	VA score	**-.01**	**.37***	.06	.06
Teacher quality	VA score	-.03	-.03	-.09	-.09
Teacher quality	Student perception	.12*	.12*	.01	.01
Managing student conduct	VA score	.14	.14	.13	.13
Instructional practice	VA score	.19	.19	.23*	.23*
Teaching workload	VA score	-.01	-.01	-.01	-.01
Instruction resources	VA score	-.07	-.07	-.03	-.03
Classroom autonomy	VA score	-.11	-.11	-.20	-.20
Managing student conduct	Student perception	**.05**	**.26***	**-.06**	**.34***
Instructional	Student	-.12	-.12	.06	.06

practice	perception				
Teaching workload	Student perception	-.19	-.19	-.17	-.17
Instruction resources	Student perception	.12	.12	.11	.11
Classroom autonomy	Student perception	.13	.13	.10	.10

*$p<0.05$

Table 4.27 showed the structural path coefficients of each of the two groups. The effect of managing student conduct on student perception was different between high and low minority schools for ELA and Math teachers. Managing student conduct had a positive and significant impact on student perception in high minority schools; while in low minority schools it showed a weak and nonsignificant effect on student perception. Additionally, the effect of student perception on teacher value-added score was different between high and low minority school, only significant in English language arts. Student perception had a positive and significant impact on teacher value-added score in high minority schools; while in low minority schools it was very weakly associated with teacher value-added score in English language arts.

Percent of Free and Reduced Price Meals Students

Studies on teacher effectiveness used the percentage of Free and Reduced Meal (FARM) students as a school demographic variable or a component of working conditions, to identify the poverty level of schools. These studies consistently find that schools with higher proportions of FARM students tend to have less effective teachers (Clotfelter, et al. 2007; Jackson, 2014; Sass, et al. 2012). To examine the impact of the percent of FARM students on the structural relationships, this study created two groups based on 1306 teachers with school context information: for schools in the top quartile of free and reduced price meals enrollment; and for schools in the bottom quartile of free and reduced price meals enrollment. Table 4.28 presented the results of statistical tests that compare differences in means or percentages. Comparison of means was based on Tamhane's T2, which does not assume equal variances across groups.

Table 4.28 *Descriptive Statistics for All Variables by Percent of Free and Reduced Price Meals (FARMS)*

	Low FARM		High FARM		Sig.
	Mean	(S.D.)	Mean	(S.D.)	
Teacher value-added score in ELA	.010	.167	-.002	.249	
Teacher value-added score in Math	.028	.203	.001	.333	
Teacher background quality	2.26	.954	2.16	1.234	
Student perception	4.01	.416	3.87	.394	***
Working conditions					
Managing student conduct	3.01	.685	2.76	.726	***
Instructional practice	3.15	.544	3.08	.511	
Teaching workload	2.71	.564	2.50	.616	***
Instruction resources	3.09	.628	2.97	.628	**
Classroom autonomy	2.90	.680	2.89	.652	

Note. Top and bottom quartiles of schools by the proportion of students eligible for Free and Reduced Price Meals (FARM). The top quartile is 87.76 percent or greater eligible for FARM; the bottom quartile is 48.04 percent or fewer eligible for FARM.
*$p<0.05$; ** $p<0.01$; ***$p<0.001$.

Teachers in low poverty (low FARM) schools had significantly better students' evaluation of their teaching in classroom. Teachers in high poverty schools had lower mean value-added scores in both ELA and Math, but the differences were not significant. The findings with regard to teacher perception data generally suggested that schools serving wealthier students have better working conditions. In the low poverty schools, teachers expressed more positive perceptions of support for managing student conduct, teaching workload, and instructional resources. Though schools serving more low-income students have lower average ratings on instructional practices support and classroom autonomy, the difference between high and low poverty schools was not statistically significant.

As indicated in Table 4.29, the chi-square difference between the constrained and unconstrained models was 34.00 with 17 degrees of freedom (p<.01) for ELA teachers and 42.38 with 17 degrees of freedom (p<.01) for Math teachers, indicating there was a significant difference between the unconstrained and constrained models for either ELA or math teachers. In other words, the unconstrained model fitted to the data more closely than the constrained model. These results indicated that factorial invariance did not exist. Next, I attempted to find out which factor loadings actually differed.

Table 4.29 *Results of Test of Invariance of Measurement Model for Percent of FARM Students*

Measurement Model	χ^2	df	p value	χ^2/df	RMSEA	CFI
ELA teacher model						
Model A (unconstrained)	1395.84	557	<.01	2.51	.066	.96
Model B (constrained)	1429.84	574	<.01	2.49	.066	.96
Model C (partial constrained)	1420.73	573	<.01	2.48	.065	.96
$\chi^2_{difference}$ (Model B-Model A)	34.00	17	<.01			
$\chi^2_{difference}$ (Model C-Model A)	24.89	16	>.05			
Math teacher model						
Model A (unconstrained)	1436.58	557	<.01	2.58	.072	.95
Model B (constrained)	1478.96	574	<.01	2.58	.072	.95
Model C (partial constrained)	1460.76	572	<.01	2.55	.071	.95
$\chi^2_{difference}$ (Model B-Model A)	42.38	17	<.01			
$\chi^2_{difference}$ (Model C-Model A)	24.18	15	>.05			

For ELA teacher model, modification indices suggested that allowing the factor loading for items SP3 on student perception to vary would improve the fit of the constrained model. For Math teacher model, modification indices suggested that allowing the factor loading for items TW4 on teaching workload and IR3 on instruction resource to vary would improve the fit of the constrained model. After free estimation of factor loading for these items, the partially constrained model was estimated again to compare that model to the fully unconstrained model using the chi-square difference statistic. The difference between this partially constrained model and unconstrained model for ELA teacher was 24.89 with 16 degrees of freedom and an associated *p* value .072 (>.05), and the difference between this partially constrained model and unconstrained model for math teacher was 24.18 with 15 degrees of freedom and an associated *p* value .062 (>.05). These results indicating that the partially constrained model and the unconstrained model did not differ significantly.

After free estimation of factor loading for these items, the statistical significance of the differences in path coefficients was tested. As indicated in Table 4.30, the impact of the percentage of FARM students had a significant impact on the structural path coefficients among variables for ELA and Math teachers. The chi-square difference between the constrained and unconstrained models was 30.71 with 13 degrees of freedom (p<.01) for ELA teachers; and 22.65 with 13 degrees of freedom (p<.05) for Math teachers. This result indicated that there was a difference between the unconstrained and constrained structural models for ELA and Math teachers. Then, an attempt was made to find out which path coefficient actually differed. To do that, modification indices were examined to assess which paths should be estimated separately for the different groups.

Table 4.30 *Results of Test of Invariance of Structural Model for Percent of FARM Students*

Structural Model	χ^2	*df*	*p* value	χ^2/df	RMSEA	CFI
ELA teacher model						
Model A (unconstrained)	1364.56	547	<.01	2.49	.066	.96
Model B (constrained)	1395.27	560	<.01	2.49	.066	.96
Model C (partial constrained)	1378.51	558	<.01	2.47		

$\chi^2_{difference}$ (Model B-Model A)	30.71	13	<.01			
$\chi^2_{difference}$ (Model C-Model A)	13.95	11	>.05			
Math teacher model						
Model A (unconstrained)	1410.59	546	<.01	2.58	.072	.94
Model B (constrained)	1433.24	559	<.01	2.56	.072	.94
Model C (partial constrained)	1428.63	558	<.01	2.56	.071	.94
$\chi^2_{difference}$ (Model B-Model A)	22.65	13	<.05			
$\chi^2_{difference}$ (Model C-Model A)	18.04	12	>.05			

For ELA teacher model, modification indices suggested that allowing the path from student perception to teacher value-added score and the path from managing student conduct to student perception, to vary would improve the fit of the constrained model. For Math teacher model, modification indices suggested that allowing the path from student perception to teacher value-added score to vary would improve the fit of the constrained model. After free estimation of the coefficients for these paths, the partially constrained model was estimated again to compare that model to the fully unconstrained model using the chi-square difference statistic. The difference between this partially constrained model and unconstrained model for ELA teacher was 13.95 with 11 degrees of freedom and an associated *p* value .236 (>.05). The difference between this partially constrained model and unconstrained model for Math teacher was 18.04 with 12 degrees of freedom and an associated *p* value .115 (>.05), indicating that the partially constrained model and the unconstrained model did not differ significantly. This result suggested that percentage of FARM students in schools where teachers taught made a significant difference in several associations among the variables of the study.

Table 4.31 *Structural Path Coefficients between Low and High Percent of FARM Students*

Paths		Standardized path coefficients			
		ELA teachers		**Math teachers**	
From	To	Low FARM	High FARM	Low FARM	High FARM
Student perception	VA score	**.02**	**.30***	**-.02**	**.21***
Teacher quality	VA score	-.04	-.03	-.06	-.06
Teacher quality	Student perception	.06	.07	-.01	-.01
Managing student conduct	VA score	.09	.09	.17	.17
Instructional practice	VA score	.10	.10	.16	.16
Teaching workload	VA score	.16	.16	.11	.11
Instruction resources	VA score	.00	.00	-.07	-.07
Classroom autonomy	VA score	-.09	-.09	-.20*	-.20*
Managing student conduct	Student perception	.13	.13	.19*	.19*
Instructional practice	Student perception	.06	.06	-.07	-.07
Teaching workload	Student perception	-.18	-.18	-.22*	-.22*
Instruction resources	Student perception	-.03	-.03	.02	.02
Classroom autonomy	Student perception	**-.03**	**.15**	.09	.09

*p<0.05

Table 4.31 showed the structural path coefficients of each of the two groups. The effects of student perception on teacher value-added score were different between high and low FARM schools for ELA and Math teachers. Student perception had a positive and significant impact on teacher value-added score in high FARM schools; while in low FARM schools it showed weak and no effect on teacher value-added score. Although the effect of classroom autonomy on student perception was significantly different between high and low FARM school, this effect of classroom autonomy on student perception was still not significant in high FARM schools.

CHAPTER 5

DISCUSSION

The purpose of this study was to identify the effects of working conditions on teacher effectiveness. This study used structural equation modeling (SEM) to examine the relationships among working conditions, teacher effectiveness, teacher background quality, and school context. In this chapter, the findings are summarized and the implications are discussed. Also, limitations of this research and recommendations for future research are included in the chapter.

Summary of Findings

Teacher effectiveness is one of the most critical issues facing the public education system in the United States. It directly influences students' learning and is the most important school-based factor in improving students' academic achievement. In recent years, researchers have consistently shown that effective teachers are distributed very unevenly among schools, especially to the clear disadvantage of high-need schools (Clotfelter, Ladd, & Vigdor, 2011; Lankford, Loeb, & Wyckoff, 2002; Sass, et al. 2012). Many attempts were made to establish an equitable distribution of effective teachers among schools through national or state educational policies (Race to the Top Progress, 2013). Although recruiting knowledgeable and skilled teachers is important, it is insufficient for schools to ensure effective teaching performance (Berry, Daughtrey, & Wieder, 2009; 2010). Good teachers need a workplace that promotes their efforts in a variety of ways to sustain their effective teaching and doing their best work with students. When the schools provide a series of supports for instruction as well as good working environment, teachers are more likely to stay and be effective in their work (Johnson et al., 2012; Loeb et al., 2005). Forecasting which of the working conditions is most likely to improve teacher effectiveness requires a better understanding of the relative importance of multiple aspects of instruction-classroom supports and how they affect teacher effectiveness as measured by teacher value-added score and student perception about teaching.

To better understand the impacts of school working conditions on teacher effectiveness, this study aimed to identify the effects of overall as well as each aspect of working conditions, teacher background quality, and the effect of high-need school context on teachers' value-added scores and student perception about teaching. Furthermore, it has examined whether all constructs and their effects are different between high-need and low-need schools.

Previous literature has found that the teachers' background qualities (e.g. teaching experience, teacher credentials, certification status, etc.) are factors to teacher effectiveness (Clotfelter et al., 2007; Darling-Hammond, 2006; Darling-Hammond & Sykes, 2003). Although findings from studies about the effects of each aspect of working conditions have been consistent, showing that schools with better instructional support, more appropriate teaching workload, better school resources, and more autonomy in classroom are more likely to have effective teaching (e.g. Johnson, 1990; McLaughlin & Talbert, 2001; Rosenholtz, 1989), often the focus has been on more physical and economic working conditions such as buildings, class size, and salaries. Little empirical work has investigated the effects of instruction and classroom related working conditions to improve teacher effectiveness, specifically improvement in teachers' value-added scores and student evaluation of teaching. Moreover, many studies have consistently shown that poor work environments and conditions that matter to effective teaching are more common in high-need schools that are attended by minority, limited English proficiency and low-income students (Berry, Smylie, & Fuller, 2008; Johnson, Kraft, & Papay, 2012; Ladd, 2011). This inequality in the distribution of effective teachers among schools is to the clear disadvantage of these high-need students (Clotfelter, Ladd, & Vigdor, 2011; Lankford, Loeb, & Wyckoff, 2002).

The current study presented a complex model of teacher effectiveness as measured by improvement in student achievement and student evaluation of teaching. These two measures are correlated, showing the criterion related validity of the student perceptions of teaching as a measure of teaching effectiveness. The model also included teacher background qualities. A two-step structural equation modeling approach was used to estimate the hypothesized model. In the first step, the measurement model was tested to examine how the latent variables were measured by sets of items. Results showed that the latent variables: student perception about teaching, teacher background quality, and six school working conditions (i.e. instructional support, teaching workload, instruction resources, classroom autonomy, and managing student conduct) were adequately reliable and valid. In the second step, the structural model was tested to estimate the path coefficients in the hypothesized model and examine how the model fit with the data.

Results showed that teacher background quality and school working conditions influenced teacher effectiveness in different ways. First, teacher background qualities, teaching experience and education, were found to positively affect student perception about teaching; and improved teacher value-added score indirectly through student perception about teaching. Furthermore, the study found the classroom-instruction level supports to be significant factors that led to teachers' effectiveness in increasing student achievement gains. In a positive working environment of support for classroom instruction, students' evaluation of teaching was also higher. Specifically, the results showed that among the working conditions, instructional practice support had a significant direct impact on teachers' value-added scores in both English language arts and mathematics compared to other working conditions. Support for managing student conduct and classroom autonomy had a weak direct effect but significant indirect effects on teachers' value-added scores. The same two working conditions showed significant impacts on students' perceptions about teaching. All the six aspects of working conditions were found to be significantly worse as perceived by teachers in high-need schools than in low-need schools. Student perception about teaching was found to be significantly worse in high-need schools than those schools serving fewer minority or students from low-incoming families, but teacher value-added score in either English language arts or mathematics were not significantly different between high-need and low-need schools. Finally, the effects of support for managing student conduct and classroom autonomy on student perception about teaching, and the relationship between student perception of teaching and teacher value-added score were also found to vary by the high-need school context. Despite important identified relationships, the amount of explained variance in the teachers' value added scores is low. There could be various reasons for it such as variables not included in the model but one reason seems to be the small overall covariation among variables in the model.

Discussion of Findings

In the following sections, the research findings are discussed in four sections: (1) the relationship between teacher value-added score and student perceptions about teaching; (2) the effects of teacher background quality; (3) the effects of each working condition; and (4) the differences in these relationships between high-need and low-need schools.

The Relationship between Teacher Value-added Score and Student Perceptions of Teaching

Results of this study identified a significant effect of students' perceptions of teaching on teachers' value-added scores, in both English language arts (ELA) and mathematics scores. This finding confirmed earlier literature that has shown a strong association between these two measures of teacher effectiveness (e.g. Bill & Melinda Gates Foundation, 2012a; Hanover Research, 2013; Raudenbush, 2013; Wilkerson, et al. 2000). Teachers whose students are satisfied with their teaching practice are more likely to gain higher value-added scores. Value added scores reflect gains in student learning. In the model, the effect of student perception about teaching on teacher value-added score was relatively strong across multiple structural models and remained consistently significant even after other working conditions were added in the models. This finding supported the argument that student-perceived rating of the teaching is a reliable quantitative measure of effective teaching and thus can successfully predict the students' achievement gain scores. In this study, students' perceptions about teaching focused on the teacher behaviors that promote student learning. In other words, this finding suggests that when students feel teachers care about their learning, are willing to work with students, and make efforts to clarify the instruction, students are likely to learn more and have higher achievement scores.

Effects of Teacher Background Quality

In the study, teacher background quality was measured by a composite variable that combines teaching experience and teacher's advanced education degree. Based on the model, the teacher background quality was found positively related to student perception about teaching. This finding was consistent with the results of some earlier studies that showed teacher background qualities are factors in teacher effectiveness (e.g. Bill & Melinda Gates Foundation, 2012a; Clotfelter, Ladd, & Vigdor, 2007; Darling-Hammond, 2006; Rivkin, Hanushek, & Kain, 2005). Findings from this study supported the notion and indicated that a teacher who has more teaching experience and/or an advanced education degree (Master degree or above) is more likely to have students who are satisfied with his/her teaching. It is noteworthy that this effect of teacher quality was only significant for ELA teachers and not for mathematics teachers. This is likely to be a technical issue resulting from sample quality, because in the sample, math teachers presented smaller variance in teaching experience and education degree level. The average education level was higher for math teachers compared to ELA teachers.

Results did not show significant direct effects of teacher background quality to teacher value-added scores in either ELA or math. Students' perceptions of teaching were more positive for more educated and experienced teachers, yet there was no statistically significant difference in value-added scores of more qualified and less qualified teachers. One possible reason is that the variance of the value-added score may be due to other unobserved teacher background qualities, such as teachers' professional experiences. Not all teachers, with the same number of teaching years in the classroom or education level, have the same teaching experiences or professional development experiences. This is consistent with some early findings that showed small or no effect of teachers' background quality such as education level and teaching experience, on student outcomes (Aaronson, Barrow, & Sander, 2007; Rivkin et al., 2005). For example, Murnane & Steele, (2007) found that teaching experience after the initial three years does not result in improved student test scores. This finding suggested that more educated and more experienced teachers teach better so that the student perceptions about teaching are more positive and affect their achievement positively.

Effects of Working Conditions

Instructional Practice Support

Instructional support includes the use of assessment data, instructional coaching and professional learning opportunities to improve and inform instructional practice. Instructional support has been operationalized similarly by other researchers, focusing on classroom level teaching support, while some researchers have also included measures of collegiality and collaborative work in instructional support (Johnson, Kraft & Papay, 2012). Prior research indicated that instructional practice supports allow teachers to establish and communicate clear learning goals to students. Average student achievement growth is higher in schools with better instructional practice supports (Brophy, 1996; Thomas & Green, 2015).

This study looked specifically at how the instructional practice supports relate to teachers' value-added scores and the students' perception about teaching quality. Consistent with what prior work suggests, the instructional practice supports had a direct positive effect on teacher value-added score. Teachers' value-added scores were higher in schools where teachers perceive more opportunities for improvement of their instruction. In schools with high instructional support teachers' learning goals are clear and are explicitly communicated to the students. The teachers in such schools are able to establish the learning goals for students, which in turn create a positive learning environment.

When teachers have access to assessment data, are provided opportunities for professional development and are supported in their efforts to provide better instruction, they are likely to be more effective in the classroom. Such pedagogical coaching and instructional support will lead to better student learning and higher student achievement gains. More specifically, in more instructionally supportive environment school administrators provide timely assessment data to teachers so they can use it to inform their instruction, teachers work in professional learning communities to develop and implement innovative instructional practices. In such schools teachers are able to try different approaches to improve learning. These instructional supports (i.e. coaching, access to professional development, use of data etc.) lead to better teaching and higher student learning. The direct significant effect of instructional supports is an important finding based on a rigorous methodology that confirms the earlier research and brings strong empirical evidence to the importance of instruction related continuous support to improve teaching effectiveness. This is the most important evidence based finding of the study.

Compared to other working conditions, the effect of instructional support on teacher value-added score was largest in both English language arts and mathematics. And this effect was significant whether or not the other working conditions were controlled in the model. Of all the working conditions, instructional practice supports stood out as the strongest factor in improving teacher effectiveness. The reason could be because schools with strong instructional support provide teaching related information to the teachers. Such environment conveys clear learning goals to teachers and students and helps teachers in monitoring student progress, providing feedback, encouraging teacher learning.

While a small positive but significant effect of instructional practice supports on student perceptions of teaching was found when it is the only working condition in the model, this effect disappears when controlling for other working conditions. Teachers who worked in schools with more supports in developing and informing their instruction obtained higher teaching evaluation from the students, however the effect of instructional support was not statistically significant on the students' perception of teaching. This effect was probably diminished when other working conditions were controlled in the model.

Teaching Workload

Teaching workload refers to the available time for teachers to plan, provide instruction, and eliminate barriers to maximize instructional time during the school day

(Ladd, 2011). This study considered the role of teaching workload as one of classroom and instruction related working conditions in teacher effectiveness. Whereas prior research indicated that more time for planning and collaboration were correlated with teachers' plans to remain in teaching, and in turn related to improved student achievement (Berry & Fuller, 2007; Louis, 1996), this study looked specifically at how teaching workload related to teachers' value-added scores and their students' perception about teaching.

Earlier studies considered teaching workload as important to teacher turnover and satisfaction (Ingersoll, 2001; Johnson, Kraft & Papay, 2012; Ladd, 2009), yet its correlation to students' learning outcome was weak (Berry & Fuller, 2007; Rice, 2009). The findings in this study showed that teaching workload did not have a significant effect on teachers' value added score, when other working conditions were controlled for. When other conditions were not in the model teachers' workload had a significant direct effect on teachers' value added score in math group only. However, no significant direct effect of workload was found on teacher value-added score in English language arts. To some degree, the results indicated that teachers were more likely to improve students' achievement gains in mathematics when they work in schools with more appropriate teaching workload. Appropriate workload seems to be more salient factor for the math teachers. It might be because math teachers usually need a larger amount of time in instruction planning and collaboration with their colleagues (Baker et al., 2004). Math teachers in schools that provide better protection from duties or routine works that interfere with the instruction are able to have sufficient instructional time.

This study showed a negative effect of appropriate teaching load on student perceptions. It seems like it is due to the correlations among appropriate teaching load and other instructional support conditions. It is noteworthy that such negative effect was detected as significant only when other working conditions were controlled in the model. A possible explanation is the correlations and interactions between teaching workload and other working conditions, resulting in the change from positive to negative effect, when other correlated factors are in the model. Teaching workload was found positively correlated to student conduct management and instruction resources. The effect of teaching workload on teacher value-added score was also weakest compared to the other working conditions. Although many studies asserted the importance of enough time in teacher's learning and satisfaction (Ware & Kitsantas, 2007; Watkins, 2005), recent research indicated that more time cannot guarantee teachers' growth and effectiveness (Coburn & Russell, 2008). For example, Coburn and Russell (2008) concluded that more time that created large opportunities for teachers to collaborate with other teachers might have limited impact on teacher effectiveness if multiple priorities compete for teachers' time and attention.

Instruction Resources

Instruction resources allow for the availability of instructional materials, technology, communication, and other instruction related resources to teachers. However, no significant effects were found from instruction resource to any of the two teacher effectiveness constructs: teachers' value-added scores and students' perception about teaching. The finding that teachers' effectiveness appeared to be unrelated to perceived instructional resources is perhaps unsurprising, given the overall results of the review of the literature on school resources (Johnson, Kraft, & Papay, 2012). Although some prior research indicated that school facility and resources were related to student achievement

(Chan & Petrie, 2000; O'Neil & Oates, 2001), the findings on the impact of resources on achievement were not consistent. For example, some studies have shown that instructional resources raised achievement in math and literacy (Giangreco & Broer, 2007; Lane, 2007), while some others have shown the opposite effect (Gerber, et al., 2001). The result of the weak effect of instruction resources on teacher effectiveness was consistent with early studies that found weak or even no effects of instruction resources on students' learning by examining a broad array of working conditions (Johnson, Kraft, & Papay, 2012). For example, as Johnson, Kraft, & Papay (2012) concluded that the conditions of school resources matter to teachers but not significantly correlated to their students' performance.

Classroom Autonomy

Classroom autonomy indicates how much teachers have control over multiple areas of planning and instruction activities in their role as teachers. The current study considered the role of classroom autonomy as an important classroom-instruction related working condition in teacher effectiveness. Numerous studies have pointed to the importance of classroom autonomy in terms of teacher retention (Guarino, Santibañez, & Daley 2006; Ingersoll & May, 2012) and student achievement (Berry, 2007; Leithwood, 2006). In this study, no significant direct effects were found from classroom autonomy to teacher effectiveness measured as teachers' value-added scores. However, slight but significant indirect relationship of classroom autonomy to teachers' value-added scores was discovered through students' perception about teaching for math teachers only. But significant direct effect was found from classroom autonomy to student perceptions of teaching for both ELA and Math teachers, which means when teachers exercise classroom autonomy and have appropriate control over material and pedagogy, they deliver better instruction according to their students. This effect was significant whether or not the other working conditions were controlled in the model. Thus, classroom autonomy had an indirect effect on teacher value-added score mediated through the students' perceived teaching in both subjects. When schools allow more autonomy and a bigger role for teachers in classroom activities, it encourages teachers to be more innovative, and try different methods to teach better. Such autonomy translates in better teaching and thus leading to higher gains in students' achievement scores. The classroom autonomy promotes effective teaching and instruction through better interactions with diverse students in the classroom (Leithwood, 2006; Nooruddin & Baig, 2014). Given that teachers have greater control over classroom activities, the students are in a better learning environment with more caring and support for their learning. Innovative pedagogy and creative ways of engaging students in learning due to greater autonomy and professional control of teachers result in higher achievement.

Support for Managing Student Conduct

Managing student conduct is in the realm of the school policies and practices designed to address student conduct issues that ensure a safe classroom environment for teachers. Prior work indicates that behavior management in classroom, which includes managing disruptive behavior among students, was significantly correlated to student achievement (Oday, 1984; Short, 1987). This study specifically examined how the better management of student conduct influenced teachers' value-added scores and their students' perceptions about teaching. Managing student conduct had direct positive effects on teacher value-added score for ELA teachers. In addition, slight but significant indirect relationship of managing student conduct to teachers' value-added scores was discovered through students' perception about teaching. Similar to classroom autonomy, significant and direct effects were found from student conduct management to student perception about teaching in both English language arts and mathematics. In schools where teachers perceived higher support in managing student conduct, they were rated higher on teaching effectiveness by their students. These effects were significant whether or not the other working conditions were controlled in the model. Consistent with what prior work suggests, teacher effectiveness and students' achievement were higher in schools where teachers perceive more support for student conduct management. In schools where rules and procedures about student conduct are implemented fairly and consistently, the policies regarding student conduct are clearly communicated, teachers can better maintain discipline and create a safer and respectful environment in the classroom. The teachers in such schools are able to effectively interact with students and maintain consistent standards of behavior, which in turn creates a positive learning environment and enhances learning outcomes.

In addition, support for managing student conduct indirectly affected teacher value-added scores through students' perceptions of teaching. These results suggest that students' perceptions about teaching mediate the link between managing student conduct and teacher value-added score. It is possible that better support in schools for student conduct management encourages more effective teaching, and in turn increases the students' achievement gains in mathematics. This result confirmed the prior findings that support for student conduct and nurturing positive interactions with students in the classroom had the potential to improve instructional outcomes (Hamilton et al., 2009).

Although the results were very similar for math and ELA teachers, there were some differences. These differences were that math teachers have higher mean scores and larger variance of value-added score than ELA teachers; and higher percent of math teachers also had advanced education degree, on average. This suggests that in the sample of the study math teachers are slightly more effective than ELA teachers. Math teachers' background qualities incorporating advanced education degree and experience were also found more related to effective teaching. In addition, the effect of managing student conduct was stronger on value-added score for ELA teachers compared to math teachers. The support for managing student conduct seems to be more salient factor for the ELA teachers.

Differences between High-need and Low-need Schools

High-need schools refer to schools that are attended by more minority (non-white) or low-income (eligible for free and reduced meal (FARM)) students, who rely most on the school for their learning. Several studies have pointed out that the inequality in the distribution of effective teachers among schools is to the clear disadvantage of high-need schools (Clotfelter, Ladd, & Vigdor, 2011; Lankford, Loeb, & Wyckoff, 2002). However, this study found that although the average teachers' value-added scores were lower in high-need schools serving more minority or FARM students, but these differences were not statistically significant. This finding indicates that after controlling for the student and classroom background variables, students' achievement gain scores were not significantly different between high-need and low-need schools. One possible explanation for the non-significant difference in teacher value-added scores could be that the sample consisted predominantly of minority (>70%) and poor students (>65%), this data limitation defuses the difference between schools that are attended by more and less minority/poor students. Although there was no significant difference in the value-added scores in the high-need versus low-need schools, there was a significant difference in the student evaluation of teaching in the two groups of schools. Significantly lower student ratings of teaching were found in high-need schools compared to the more advantaged schools. That means, teachers who work in high-need schools that are attended by minority or FARM student had lower student evaluation of teaching. This finding is consistent with several studies about the shortage of effective teachers in schools serving primarily disadvantaged and minority students versus in schools with more advantaged students (Clotfelter et al., 2005; Lankford et al., 2002). According to the perceptions of students, teachers were less caring about student learning, and they often did not provide clarification or consolidation in instruction at high-need schools.

While in exploring the teacher effectiveness gap between high-need and low-need schools, most studies investigated the difference in teacher characteristics or background qualities, but did not investigate how the working conditions contribute to the teacher effectiveness gap. Recent large-scale quantitative studies have provided evidence that poor work environments and conditions that matter to teachers are more common in high-need schools (Berry, Smylie, & Fuller, 2008; Johnson, Kraft, & Papay, 2012; Ladd, 2011). Consistent with what prior studies suggest, the descriptive statistics and t-test comparison results indicated that teachers in high-need schools often perceived lower support for management of student conduct, inadequate support for instructional practice, inappropriate instruction workload, lack of instruction resources, and less autonomy in classroom activities. These significant differences in the two groups are the likely factors for differences in student evaluation of teaching which is strongly related to value-added gains.

Furthermore, the SEM invariance testing results indicated that the effects of working conditions on teacher effectiveness were different in the two groups. Particularly, certain aspects of school working conditions, including support for managing student conduct and classroom autonomy, mattered more in high-need schools serving disadvantaged students. Support for managing student conduct showed a significantly larger effect on student perception about teaching in high-need schools than in more advantaged schools. This finding is consistent with other research that suggests that the quality of working conditions may be especially impactful in high-need schools, in that working conditions have a stronger impact on teacher effectiveness in high-need schools (Grissom, 2011). In addition, student perception about teaching was found more strongly related to teachers' value-added score in high-need schools serving more minority or poor students. The relationship between these two measures of teacher effectiveness was weaker and even not significant (at the edge of the significance) in low-need schools. This finding indicated that students' evaluation of teaching better predicted teachers' value-added scores in high-need schools than in more advantaged schools.

Implications

The findings of the study have both theoretical and practical implications. The identified relationships among teacher effectiveness, school working conditions, and teacher background qualities have been generally studied separately without control variables, and the mediating effects are often overlooked, thus the effects are often biased in extant literature. This work fills a gap in current knowledge and provides convincing evidence through rigorous methods for the important relationships in the following five areas: (1) the study provides empirical evidence to the important role of student perceptions about teaching in measuring teacher effectiveness; (2) the findings support the important role of classroom and instruction related working conditions in teacher effectiveness in both subjects; (3) the study confirms the impact of instructional support on teacher value-added scores, comparing it with other working conditions; (4) the findings suggest that support for managing student conduct and classroom autonomy are other influential working condition factors that affect teacher value-added scores. These factors have indirect effects on teacher value-added scores through the mediating effects of student satisfaction with teaching; and (5) the study further confirms the effect of teacher background quality on student perceptions of teaching and indirectly on value-added scores in ELA group.

One important finding from the current study is that students' satisfaction with teaching can predict their learning outcome, as can be seen by the significant effects of student perceptions about teaching on teacher value-added scores in both ELA and mathematics. The findings of the study provide empirical evidence that can be used by policy makers or educational researchers in measuring or evaluating teacher effectiveness. Using a single teacher attribute or characteristic is not adequate to define an effective teacher (Papanastasiou, 1999). The study confirms that student rating of teaching can be used as a complement to other tools, such as classroom observations or measures of student achievement gains in the evaluation of effective teachers. Data from student survey captured the impressions of many individuals who've spent many hours with the teachers; it provides many details similar to classroom observations. Moreover, it is important to note that student' perceptions of teaching predicted students' achievement gain scores more precisely in high-need schools serving minority or poor students.

The study provides some ideas for improving the working conditions of the teachers. Better support for instruction, timely data access, opportunities for professional development, coaching and mentoring of teachers for trying innovative pedagogy are all likely to lead to more effective teaching. Teachers need professional learning communities for continuous improvement in their teaching practices. Secondly, most previous studies that have focused on teacher effectiveness have not taken into account the effect of support for managing student conduct and classroom autonomy (Berry, Daughtrey, & Wieder, 2010; Berry, Smylie, & Fuller, 2008). The present study highlights that better support for student conduct management and classroom autonomy is likely to lead to more effective teaching, and in turn to increase in student achievement gain scores. These classroom management and safe environment conditions are salient, as then more attention can be paid to the classroom activities, interaction with students, to achieve high quality and effective teaching, especially in high-need schools. Teachers need support in effectively managing student behaviors and classroom autonomy to become better teachers. Overall, the findings of the study significantly contribute to a better understanding of the effects of working environment and how these are related to teacher performance. The study considered a complex framework for the study of work environment factors and how these are linked to two critical measures of effective teaching: student achievement and student satisfaction with teaching.

As a result, the findings of the study provide critical evidence that can be used by policy makers or educational researchers to promote teachers' performance. The identified associations between working conditions and teacher performance have important implications for policy makers and practitioners and provide ideas to increase teacher effectiveness by altering their working conditions. Many factors in the work environment that enhance teaching effectiveness can be changed such as better support for instruction and clear policies on fair and consistent discipline. In fact, these factors can be modified more easily by effective leadership than some of the resources such better technology or physical space. Interestingly, instruction resources were not found to be significant factors in enhancing teacher effectiveness. The findings emphasize the importance of working conditions which are consistent with previous work regarding the relationships between working conditions and teacher satisfaction (Johnson, Kraft & Papay, 2012; Ladd, 2009), turnover (Ingersoll, 2001), and student achievement (Bryk and Schneider, 2002; Jackson, 2014). The results of the study support the importance of classroom-instruction related working conditions in teachers' effectiveness: instructional practice support, support for managing student conduct, and classroom autonomy. It is surely important to have adequate resources, and sufficient time for preparation, but if teachers are to achieve success with their students, they should be able to count on the on-going instructional support, control of classroom activities, and management of student behaviors. In addressing teacher effectiveness, many U.S. states and districts have implemented policies to enhance physical working conditions, professional development opportunities, and instructional programs. In the development of such programs, policy makers should consider the relative importance of these factors as they relate to student achievement and students' satisfaction with teaching.

As spurred by Race to the Top, the states are increasingly focused on developing innovations and reforms to reduce the disparities of effective teachers among schools serving diverse students. This study found the differences in students' evaluation of teaching and working conditions' effects between high-need and low-need schools. The problem of disparity is a complex one and requires many innovative ideas to reduce the gap in effective teaching. Replacing current teachers with new ones with more advanced education degrees or with more experience may not be sufficient or possible. The study suggests investing continuously in teacher development. The study found evidence that the support for managing student conduct and classroom autonomy for teachers is more impactful in high-need schools. Thus, the findings of the study have implications for decision makers and school administrators. The study suggests clear direction for developing and implementing school policies to support better teaching.

Contribution of the Study

Conceptual Contribution

The main purpose of this study was to uncover the relationships between teaching effectiveness and teachers' working conditions that directly related to instruction and classroom management from multiple perspectives. The main contribution of this study is an expanded framework to understand and examine the relationship between teaching effectiveness and teachers' working conditions. The study incorporated the overall and each working condition separately in the model to explore the differences in their impact on teacher effectiveness. Moreover, the study included the measures of teacher effectiveness from two domains: student achievement outcome and students' satisfaction with teaching practices. The model in the current study presented a complex view of teacher effectiveness by hypothesizing the relationships between the two domains of teacher effectiveness and working conditions. Whereas previous work focused on outcomes less directly linked to educational productivity, such as teacher turnover, or on school-level academic achievement, multiple measures of teacher effectiveness such as average student achievement gain scores and students' perception about classroom teaching were used in the study. The results of the study present a conceptual framework of how teaching is not a static trait but is a dynamic construct which interacts with the contextual factors and changes in response to the conditions of teaching. Thus, the working conditions, school context, and teacher background quality together lead to more effective teaching. In addition, working conditions and teacher effectiveness constructs were compared and the differences in their relationships were tested between high-need and low-need schools. The findings of interactions between working conditions and school context suggested how school context interacts with practices and policies, and what conditions are required to support school improvement efforts. The study also provides ideas for future research.

Methodological Contribution

In terms of methodology, this study provides more accurate estimates of the relationships among factors by using a structural equation modeling approach rather than the regression type analyses. The use of SEM not only enabled the specification of relations between observed variables and latent constructs, but also the relationship among latent constructs. The mediating effects of students' perception of teaching and the correlations between exogenous variables could not be assessed without the use of this methodology. The SEM approach considers both measurement error and interrelationships among factors in a comprehensive analysis. SEM provides the fit indices for assessing how well the model fits the data. Along with using the cross-validation method in this study, it examined the model robustness and helped in establishing the validity of the final models. Moreover, the application of invariance testing of SEM allows for more stringent comparison of these relationships between high-need and low-need schools. Earlier studies have made comparisons, using separate groups but simultaneous group differences analyses provide more precise estimates.

Limitations of the Study

Although this study makes theoretical and methodological contributions to research, two major limitations exist. One limitation of the study pertains to the reliability of self-reported data. Such data may contain measurement error that may not accurately represent what happened in fact. Instead of using the teacher perceived working conditions, the observed data on each working condition construct could be used in future studies so a better understanding of how the working conditions influence teacher effectiveness can be explored. In addition, the measure of teacher effectiveness used—students' average standardized test gain scores—is an attempt to capture teaching outcome and thus represents only a portion of the curriculum and successful teaching. A teacher's value-added score is not necessarily indicative of good teaching practice and an outcome based on test scores alone cannot capture the full complexity of high quality teaching or the full aim of education (Fenstermacher & Richardson, 2005). Moreover, in terms of another important measure of teacher effectiveness—students' perception about teaching practice, scholars have argued that it too cannot completely capture teachers' performance in the curriculum and instruction (Balch, 2012; Goe, 2008). Students' perception about teaching practices only accounts for teachers' performance related to students and from students' point of view. Although students' perception is an important source of information about the classroom, it may contain bias to some degree. Despite limitation of each measure separately the inclusion of both measures provided a more comprehensive view of teacher effectiveness. Another data limitation was that only two observable characteristics of teachers (teaching experience and education degree level) were available in the data. Information on certification and license was not included in the data.

Another limitation of this study is that causality may only be tentatively inferred due to cross-sectional data exploration. Although this study provides a description of working conditions that are related to teacher effectiveness, it is not intended to suggest strong causality. Students and teachers are not randomly assigned to schools and classrooms; no statistical model can fully address this lack of randomization, and so estimates of the effects based on these models can only be tentatively interpreted as causal (Rothstein, 2010). The relationship between working conditions and teacher effectiveness may exist for several reasons. For example, it is not clear whether instructional practice support makes teachers more effective, or whether more effective teachers are more inclined to seek support and perceive higher support. Better schools with strong instructional support may be able to attract more teaching candidates; this may provide more choice for hiring good teachers. It is also possible that effective teachers are particularly drawn to schools with supportive environments, or that schools with such conditions are more likely to retain their effective teachers. Despite the strengths of the conceptual model presented in the current study, only tentative causal inferences can be drawn based on cross-sectional data.

In addition, the use of MET data in developing a proxy for teacher effectiveness limits the generalizability of findings to teachers in elementary and middle public schools in the U.S. with similar characteristics. Despite these limitations the study contributes to better understanding of teacher effectiveness and working conditions, and has implications for practice.

Directions for Future Research

This study points to several potential areas for future research. Given that these findings are based on non-experimental data, further research could be conducted to determine whether policies that promote classroom management or instructional support to improve teacher practice can enhance teacher effectiveness, using randomized studies. In addition, this study focused on the influence of classroom and instruction level working conditions, other factors could be considered in future studies. For example, the factors that consider teacher values, traditions, accountability, and community and parent involvement, etc. are not only found important to teaching practice, but also found influential in teacher turnover or teacher satisfaction (Berry, Smylie, & Fuller, 2008; Ingersoll, 2001; Johnson, 2006). These findings certainly provide a rationale for more research, based on new data. Furthermore, more demographic variables such as gender, ethnicity should be included in understanding differences in teacher effectiveness. More qualitative and mixed methods research could provide better understanding and in-depth knowledge of high-quality teaching.

Future researchers could reexamine the results of this study by using the alternative statistical methods or replicating the study with new data. Other statistical models, such as hierarchical linear modeling (HLM) technique could be used to examine the working conditions' effects for testing the replicability of results and to explain the school level variability. Future studies could also consider the non-linear effects of working conditions. For example, it is possible that the effect of instruction resources and teaching workload could be curvilinear, being positive to some extent and flattening out after that point. Future studies could incorporate both interactional and non-linear effects in models of working conditions. In addition, the teacher value-added score was estimated by using a context-adjusted 2-level hierarchical model (White & Rowan, 2014). Other models, such as fixed effects value-added model (Harris & Sass, 2006), NYC value-added model (Value-Added Research Center, 2010) could be used as alternative models to estimate teacher value-added score. Certainly the study points to the need for more research using other methods and different conceptualization of the study's constructs.

REFERENCES

2014 North Carolina Teacher Working Conditions Survey. (2014 Spring). *Design, Validity and Reliability*. Santa Cruz, CA: New Teacher Center.

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*, 95-135.

Acock, A. C. (2005). Working with missing values. *Journal of Marriage and Family, 67*(4), 1012-1028.

Adelman, N. E., Eagle, K. P. W., & Hargreaves, A. (1997). *Racing with the clock: Making time for teaching and learning in school reform.* New York: Teachers College Press.

Aleamoni, L. M. (1981). *Student ratings of instruction.* CA: Sage.: Beverly Hills.

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An Interaction-Based Approach to Enhancing Secondary School Instruction and Student Achievement. *Science, 333*(1034-1037).

Allison, P. D. (2003). Missing Data Techniques for Structural Equation Modeling. *Journal of Abnormal Psychology, 112*(4), 545-557.

Altschul, I. (2010). Parental Involvement and the Academic Achievement of Mexican American Youths: What Kinds of Involvement in Youths' Education Matter Most? *Social Work Research, 35*(3), 159-170.

Bagozzi, R. R., & Yi, Y. (1988). On the Evaluation of Structural Equation Models. *Journal of the Academy of Marketing Science, 16*(1), 74-94.

Balch, R. (2012). *The Validation of a Student Survey on Teacher Practice*: Vanderbilt University.

Ball, D. L., & Cohen, D. K. (1999). Developing practices, developing practitioners: Toward a practice-based theory of professional development. *Teaching as the learning profession: Handbook of policy and practice*. San Francisco, CA: Jossey-Bass.

Baker, D. P., Fabrega, R., Galindo, C., & Mishook, J. (2004). Instructional time and national achievement: Cross-national evidence.*Prospects: Quarterly Review of Comparative Education, 34*(3), 311–334.

Bauch, P. A., & Goldring, E. B. (2000). Teacher Work Context and Parent Involvement in Urban High Schools of Choice. *Educational Research and Evaluation, 6*(1), 1-23.

Becker, H. J., & Epstein, J. L. (1982). Parent Involvement: A Survey of Teacher Practices *The Elementary School Journal, 83*(2), 85-102.

Berry, B. (2007). *Recruiting and retaining quality teachers for high-needs schools: Insights from NBCT summits and other policy initiatives.* Hillsborough, NC: Center for Teaching Quality.

Berry, B., Daughtrey, A., & Wieder, A. (2009). *Teaching effectiveness and the conditions that matter most in high-needs schools: A policy brief*. Center for Teaching Quality. http://www.teachingquality.org/sites/default/files/Tch_effective_twc_final.pdf

Berry, B., Daughtrey, A., & Wieder, A. (2010). *Teacher Effectiveness: The Conditions that Matter Most and a Look to the Future.*: Center for Teaching Quality.

Berry, B., & Fuller, E. (2007). *Stemming the tide of teacher attrition: How working conditions influence teacher career intentions and other key outcomes in Arizona*. Hillsborough, NC: Center for Teaching Quality.

Berry, B., Smylie, M., & Fuller, E. (2008). *Understanding Teacher Working Conditions: A Review and Look to the Future.* Center for Teaching Quality.

Betts, J. R., Rueben, K. S., & Danenberg, A. (2000). *Equal Resources, Equal Outcomes? The Distribution of School Resources and Student Achievement in California*. San Francisco, California: Public Policy Institute of California.

Biancarosa, G., Bryk, A. S., & Dexter, E. R. (2010). Assessing the Value-Added Effects of Literacy Collaborative Professional Development on Student Learning. *The Elementary School Journal, 111*(1), 7-34.

Bifuh-Ambe, E. (2013). Developing successful writing teachers: Outcomes of professional development exploring teachers' perceptions of themselves as writers and writing teachers and their students' attitudes and abilities to write across the curriculum. *English Teaching: Practice and Critique, 12*(3), 137-156.

Bill & Melinda Gates Foundation. (2010a). *Learning About Teaching: Initial Findings from the Measures of Effective Teaching Project*: MET Project Policy and Practice Brief. Bill & Melinda Gates Foundation.

Bill & Melinda Gates Foundation. (2010b). *Working with Teachers to Develop Fair and Reliable Measures of Effective Teaching*: MET Project Research Paper.

Bill & Melinda Gates Foundation. (2012a). *Asking Students about Teaching: Student Perception Surveys and Their Implementation: MET Project Policy and Practice Summary*. Bill & Melinda Gates Foundation.

Bill & Melinda Gates Foundation. (2012b). *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains: MET Project Policy and Practice Summary*. Bill & Melinda Gates Foundation.

Blanc, S., Christman, J. B., Liu, R., Mitchell, C., Travers, E., & Bulkley, K. E. (2010). Learning to Learn From Data: Benchmarks and Instructional Communities. *Peabody Journal of Education, 85*(2), 205-225.

Bobbitt, S., Leich, M., Whitener, S., & Lynch, H. (1994). *Characteristics of stayers, movers, and leavers: Results from the teacher follow up survey, 1991-92*. Washington, DC: National Center for Education Statistics.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. . *Educational Researcher, 30*(8), 3-15.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2008). *Who Leaves? Teacher Attrition and Student Achievement.* NBER Working Papers 14022. National Bureau of Economic Research, Cambridge, MA.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*, 416-440.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: how teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management, 24*(1), 113-132.

Boyd, D., Lankford, H., Loeb, S., Rockoff, J. & Wyckoff, J. (2008). The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High-Poverty Schools. *Journal of Policy Analysis and Management*, 27(4), 793-818.

Braun, H., Chudowsky, N., & Koenig, J. (2010). Getting Value Out of Value-Added: Report of a Workshop. Washington, DC.: National Research Council and National Academy of Education.

Brophy, J. (1996). *Teaching: A special report reprinted by the Laboratory for Student Success.* Philadelphia, PA: The Mid-Atlantic Regional Educational Laboratory at the Temple University Centre for Research in Human Development and Education.

Bryk, A. S., & Schneider, B. (2002). *Trust in Schools: A Core Resource for Improvement.* New York: Russell Sage Foundation.

Buckley, J., Schneider, M., & Shang, Y. (2004). The effects of school facility quality on teacher retention in urban school districts. Washington, D.C.: National Clearinghouse for Educational Facilities

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming.* Mahwah, NJ: L. Erlbaum Associates.

Chan, T. C., & Petrie, G. (2000). A well designed school environment promotes brain learning. *Educational Facility Planner, 35*(3), 12-15.

Charles, M. C. (2008). *Building Classroom Discipline*. Boston, MA: Pearson/Allyn and Bacon.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The Long-term Impacts of Teachers: Teacher Value-added and Student Outcomes in Adulthood*.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2013). *Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-added Estimates*.

Choy, S. P. (1996). Teachers' Working Conditions. Washington, DC: National Center for Education Statistics.

Clark, D. (1993). *Teacher evaluation: A review of the literature with implications for educators.* California State University at Long Beach.

Clifford, M., Menon, R., Gangi, T., Condon, C., & Hornung, K. (2012). Measuring School Climate for Gauging Principal Performance: A Review of the Validity and Reliability of Publicly Accessible Measures. A Quality School Leadership Issue Brief. *American Institutes for Research*.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review, 24*(4), 377-392.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). How and Why Do Teacher Credentials Matter for Student Achievement? *NBER Working Papers 12828. National Bureau of Economic Research, Cambridge, MA.*

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., Wheeler, J. (2007). High Poverty Schools and the Distribution of Teachers and Principals. *North Carolina Law Review,* 85, 1345-1379.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2011). Teacher mobility, school segregation, and pay-based policies to level the playing field. *Education Finance and Policy, 6*(3), 399-438.

Coburn, C. and Russell, J. (2008). District Policy and Teachers' Social Networks. *Educational Evaluation and Policy Analysis*. September. 30, 3, pp 203-35.

Cohen, D. K., & Ball., D. L. (1999). Instruction, Capacity, and Improving Instruction.: CPRE Research Report.

Cohen, D. K., & Hill, H. C. (2001). *Learning policy*. New Haven, CT: Yale University Press.

Collins, A. (1990). *Transforming the assessment of teachers: Notes on a theory of assessment for the 21st century*. Paper presented at the Annual meeting of the National Catholic Education Association, Boston, MA.

Cosner, S. (2011). Teacher Learning, Instructional Considerations and Principal Communication: Lessons from a Longitudinal Study of Collaborative Data Use by Teachers. *Educational Management Administration & Leadership, 39*(5), 568-589.

Cudeck, R., & Browne, M. W. (1983). Cross-validation of Covariance Structures. *Multivariate Behavioral Research, 18*(147-167).

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1).

Darling-Hammond, L. (2006). *Powerful Teacher Education: Lessons from Exemplary Programs*: San Francisco: John Wiley and Sons, Inc. 21.

Darling-Hammond, L., & Post, L. (2000). Inequality in Teaching and Schooling: Supporting High-Quality teaching and leadership in Low-income schools.

Darling-Hammond, L., & Sykes, G. (2003). Wanted: A National Manpower Policy for Education.

Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" tell us? *Educational Researcher, 31*(9), 13-25.

Datnow, A., Park, V., & Kennedy-Lewis, B. (2012). High School Teachers' Use of Data to Inform Instruction. *Journal of Education for Students Placed at Risk, 17*(4), 247-265.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S. & Birman, B. F. (2002). Effects of Professional Development on Teachers' Instruction: Results from a Three-Year Longitudinal Study. *Educational Evaluation and Policy Analysis*, 24(2), 81-112.

Duyar, I. (2010). Relationship between school facility conditions and the delivery of instruction: Evidence from a national survey of school principals. *Journal of Facilities Management*, 8(1), 8-25.

Earthman, G. I. (2002). The effects of the conditions of school facilities on student achievement: Williams v. State of California.

Elmore, R., & Burney., D. (1997). Investing in Teacher Learning: Staff Development and Instructional Improvement in Community School District#2, New York City. Washington,DC: National Commission on Teaching and America's Future.

Epstein, J. L., Sanders, M. G., Simon, B. S., Salinas, K. C., Jansorn, N. R., & Voorhis, F. L. V. (2002). *School, Family, and Community Partnerships: Your Handbook for Action*. Thousand Oaks, CA: Corwin Press.

Fan, X. (2001). Parental Involvement and Students' Academic Achievement: A Growth Modeling Analysis. *Journal of Experimental Education, 70*(1), 27-61.

Feng, L., Figlio, D., & Sass, T. R. (2010). School Accountability and Teacher Mobility. *NBER Working Papers 16070. National Bureau of Economic Research, Inc., Cambridge, MA.*

Fenstermacher, G.D. & Richardson, V. (2005). On Making Determinations of Quality in Teaching. *Teachers College Record, 107*(1), 186–213.

Firestone, W. A., & Pennell, J. R. (1993). Teacher commitment, working conditions, and differential incentive policies. *Review of Educational Research, 63*(4), 489-525.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 286-299.

Follman, J. (1992). Secondary school students' ratings of teacher effectiveness. *The High School Journal, 75*(3), 168-178.

Follman, J. (1995). Elementary public school pupil rating of teacher effectiveness. *Child Study Journal, 25*(1), 57-78.

Fornell, C., & Larcker, D. (1981). Structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Futernick, K. (2007). *Excellence loves company: A tipping point turnaround strategy for California's low performing schools*. San Francisco: WestEd.

Gerber, S.B., Finn, J.D., Achilles, C.M. & Boyd-Zaharias, J. (2001). Teacher aides and students' academic achievement. *Educational Evaluation and Policy Analysis, 23*(2), 123-143.

Gerstle, L., & French, D. (1993). *Structuring schools for student success: A focus on Instructional improvement*: Massachusetts State Dept, of Education, Quincy.

Gettinger, M., & Guetschow, K. W. (1998). Parental involvement in schools: Parent and teacher perceptions of roles, efficacy, and opportunities. *Journal of Research & Development in Education, 32*(1), 38-52.

Giangreco, M.F. & Broer, S.M. (2007). School-based screening to determine overreliance on paraprofessionals. *Focus on Autism and Other Developmental Disabilities, 22*(3), 149-158.

Glazerman, S., Loeb, S., Goldhaber, D., Staiger, D., Raudenbush, S., & Whitehurst, G. (2010). *Evaluating Teachers: The Important Role of Value-Added*. Washington DC: Brown Center on Education Policy at Brookings.

Goddard, R., Tschannen-Moran, M., & Hoy, W. (2001). A multilevel examination of the distribution and effects of teacher trust in students and parents in urban elementary schools. *The Elementary School Journal, 102*(1), 3-17.

Goddard, Y., & Goddard, R. D. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record, 109*(4), 877-896.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis.* Washington, DC.: National Comprehensive Center for Teacher Quality.

Goe, L., Bell, C., & Little, O. (2008). *Approaches to Evaluating Teacher Effectiveness: A Research Synthesis.*National Comprehensive Center for Teacher Quality.

Goldhaber, D., & Brewer, D. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *The Journal of Human Resources, 32*(3), 505-523.

Grewal, R., Cote, A.J., & Baumgartner, H., (2004), "Multicollinearity and measurement error in structural equation models: implications for theory testing," *Marketing Science*, 23(4), 519-529.

Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). *Using student achievement data to support instructional decision making.* Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Hanover Research (Feb, 2013). Student Perception Surveys and Teacher Assessments: Washington, DC

Hanushek, E. A. (1971). Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data. *American Economic Review Papers and Proceedings, 61*(2), 280-288.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1999). *Do Higher Salaries Buy Better Teachers?* National Bureau of Economic Research, Inc.

Hanushek, E. A., & Rivkin, S. G. (2007). Pay, Working Conditions, and Teacher Quality. *Future of Children, 17*(1), 69-96.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources, 39*(2), 326-254.

Hanushek, E. A., & Rivkin, S. G. (2010). Using value-added measures of teacher quality. : National Center for Analysis of Longitudinal Data in Education Research. Urban Institute.

Harris, D. N., Sass, T. R. (2006). Value-Added Models and the Measurement of Teacher Quality. *Draft for Teacher Quality Research*.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95*(7-8), 798-812.

Heneman, H. G., Milanowski, A., Kimball, S. M., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay.* Philadelphia: Consortium for Policy Research in Education.

Henson, R.K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development,* 34 (3), 177-189.

Hill, N. E., & Tyson, D. F. (2009). Parental involvement in middle school: A meta-analytic assessment of the strategies that promote achievement. *Developmental Psychology, 45*(3), 740-763.

Holtzapple, E. (2003). Criterion-related validity evidence for a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education, 17*(3), 207-219.

Honig, M. I., Kahne, J., & McLaughlin, M. W. (2001). School-Community Connections: Strengthening Opportunity to Learn and Opportunity to Teach. In V. Richardson (Ed.), *Handbook of Research on Teaching* (pp. 998-1028). Washington, DC: AERA.

Ingersoll, R. M. (1995). *Teacher supply, teacher qualitications and teacher turnover*. Washington, DC: National Center for Education Statistics.

Ingersoll, R. M. (2001). *Teacher turnover, teacher shortages, and the organization of schools*: Seattle, WA: University of Washington, Center for the Study of Teaching and Policy.

Ingersoll, R. M. (2003). *Who controls teachers' work?: Power and accountability in America's schools.* Cambridge, MA Harvard University Press.

Ingersoll, R. M., & May, H. (2012). The Magnitude, Destinations, and Determinants of Mathematics and Science Teacher Turnover. *Educational Evaluation and Policy Analysis, 34*(4), 435-464.

Jackson, C. (2014). *Are Working Conditions Related To Teacher Effectiveness?* 39th Annual Conference of the Association for Education Finance and Policy, San Antonio, Texas.

Jackson, C. K. (2010). Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence From Teachers. *NBER Working Paper*.

Jackson, C. K., & Bruegmann, E. (2009). Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics, 1*(4), 85-108. doi: doi: 10.1257/app.1.4.85

Jeynes, W. H. (2007). The Relationship between Parental Involvement and Urban Secondary School Student Academic Achievement: A Meta-Analysis. *Urban Education, 42*(1), 82-109.

Johnson, S. M. (1990). *Teachers at Work: Achieving Success in Our Schools.* New York: Basic Books.

Johnson, S. M. (2006). *The Workplace Matters: Teacher Quality, Retention, and Effectiveness*. NEA Professional Library.

Johnson, S. M., & Birkeland, S. E. (2003). The schools that teachers choose. *Educational Leadership, 60*(8), 20-24.

Johnson, S. M., Kraft, M. A., & Papay, J. P. (2012). How Context Matters in High-Need Schools: The Effects of Teachers' Working Conditions on Their Professional Satisfaction and Their Students' Achievement. *Teachers College Record, 114*(10), 1-39.

Jöreskog, K. G., & Sörbom, D. (2006). LISREL 8.8. Chicago: Scientific Software International.

Kane, T. (2012). Capturing the dimensions of effective teaching: student achievement gains, student surveys, and classroom observations. *Education Next,* 12(4), 34.

Kane, T., & Staiger, D. (2008). Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. *NBER Working Paper 14607, National Bureau of Economic Research, Inc*.

Kendziora, K., & Osher, D. (2009). *Starting to turn schools around: The academic outcomes of the safe schools, successful students initiative*. Washington, DC: American Institutes for Research.

Kimball, S. M., White, B., Milanowski, A. T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education, 79*(4), 54-78.

Kline, R. B. (2011). *Principles and practices of structural equation modeling*. New York: Guilford Press.

Kraft, M. A., & Papay, J. P. (2014). Do supportive professional environments promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*.

Kruse, S., & Louis, K. S. (2009). *Strong Cultures: A Principal's Guide to Change*. Thousand Oaks, CA: Sage.

Kyriakides, L. (2005). Drawing from Teacher Effectiveness Research and Research into Teacher Interpersonal Behaviour to Establish a Teacher Evaluation System: A Study on the Use of Student Ratings to Evaluate Teacher Behaviour. *Journal of Classroom Interaction, 40*(2), 44-66.

Lance, C.E., Butts, M.M., & Michels, L.C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods,* 9 (2), 202-220.

Ladd, H. F. (2009). *Teachers' Perceptions of their Working Conditions: How Predictive of Policy-Relevant Outcomes?* CALDER, The Urban Institute. Retrieved from http://www.urban.org/uploadedpdf/1001440-Teachers-perceptions.pdf

Ladd, H. F. (2011). Teachers' Perceptions of Their Working Conditions: How Predictive of Planned and Actual Teacher Movement? *Educational Evaluation and Policy Analysis, 33*(2), 235-261.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis, 24*(1), 37-62.

Lane, K.L., Fletcher, T., Carter, E.W., Osdejud, C. & DeLorenzo, J. (2007). Paraprofessional-led phonological awareness training with youngsters at risk for reading and behavioral concerns. *Remedial and Special Education, 28*(5), 266-276.

LaRocque, M., Kleiman, I., & Darling, S. M. (2011). Parental Involvement: The Missing Link in School Achievement. *Preventing School Failure: Alternative Education for Children and Youth, 55*(3), 115-122.

Learning First Alliance. (2005). *A shared responsibility: Staffing all high-poverty, low-performing schools with effective teachers and administrators*. Washington, DC: Author.

Leithwood, K. (2006). *Teacher Working Conditions That Matter: Evidence for Change.* Toronto: Elementary Teachers' Federation of Ontario.

Little, J. W. (1990). The Persistence of Privacy: Autonomy and Initiative in Teachers' Professional Relations. *Teachers College Record, 91*(4), 509-536.

Little, J. W. (1993). Teachers' professional development in a climate of education reform. *Educational Evaluation and Policy Analysis, 15*(2), 129-151.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*(151-173).

Liu, X. S., & Meyer, J. P. (2005). Teachers' perceptions of their jobs: A multilevel analysis of the teacher follow-up survey for 1994-95. *Teacher College Record, 107*(5), 985-1003.

Lockwood, J. R., & McCaffrey, D. (2009). Exploring Student-Teacher Interactions in Longitudinal Achievement Data. *Education Finance and Policy, 4*(4), 439-467.

Loeb, S., Darling-Hammond, L., & Luczak, J. (2005). How Teaching Conditions Predict Teacher Turnover in California Schools. *Peabody Journal of Education, 80*(3), 44-70.

Loeb, S., Elfers, A. M., Knapp, M. S., & Plecki, M. L. (2004). *Preparation and support for teaching: A survey of working conditions of teachers.* University of Washington, Seattle: Center for Strengthening the Teaching Profession.

Loeb, S., Kalogrides, D., & Béteille, T. (2011). *Effective Schools: Teacher Hiring, Development, and Retention*. Cambridge, MA.

Loehlin, J. C. (1987). *Latent Variable Models*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Louis, K. S., Kruse, S., & Marks, H. (1996). Teachers' Professional Community in Restructuring Schools. American Educational Research Journal, 33(4), 757-798.

MacCallum, R. C., Roznowski, M., Mar, C. M., & Reith, J. V. (1994). Alternative Strategies for Cross-Validation of Covariance Structure Models. *Multivariate Behavioral Research, 29*(1), 1-32.

Mapp, K. L. (1999). *Making the Connection between Families and Schools: Why and How Parents Are Involved in Their Children's Education.* (PhD), Harvard University, Cambridge, MA.

Markley, T. (2006). Defining the Effective Teacher: Current Arguments in Education.

McGarity, J. R., & Butts, D. P. (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 and Science Teaching, 21*(1), 55-61.

McKinney, S. E., Campbell-Whately, G. D., & Kea, C. D. (2005). Managing Student Behavior in Urban Classrooms: The Role of Teacher ABC Assessments. *The Clearing House, 79*(1), 16-20. doi: 10.2307/30182100

McLaughlin, M. W., & Talbert, J. E. (2001). *Professional Communities and the Work of High School Teaching.* Chicago: University of Chicago Press.

Million, S. (1987). *Demystifying teacher evaluation: The multiple-strategies model used as an assessment device.* Paper presented at the Annual meeting of the National Council of States on in-Service Education, San Diego, CA.

Molly F. Gordon, & Karen Seashore Louis. (2009). Linking Parent and Community Involvement with Student Achievement: Comparing Principal and Teacher Perceptions of Stakeholder Influence. *American Journal of Education, 116*(1), 1-31. doi: 10.1086/605098

Muntner, M. (2008). Teacher-Student Interactions: The Key To Quality Classrooms. The University of Virginia Center for Advanced Study of Teaching and Learning (CASTL).

Murnane, R. J. (1981). Teacher mobility revisited. *Journal of Human Resources, 16*(1), 3-19.

Murnane, R. J., & Steele, J. L. (2007). What is the problem? The challenge of providing effective teachers for all children. *The Future of Children, 17*(1), 15-43.

Nooruddin, S., & Baig, S. (2014). Student behavior management: School leader's role in the eyes of the teachers and students. *International Journal of Whole Schooling, 11*(1), 19-39.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? . *Educational Evaluation and Policy Analysis, 26*, 237-257.

O'Neil, D. J., & Oates, A. D. (2001). The impact of school facilities on student achievement, behavior, attendance, and teacher turnover rate in Central Texas Middle Schools. *Educational Facility Planner, 36*(3), 14-22.

Obenchain, K. M., & Taylor, S. S. (2005). Behavior management: Making it work in middle and secondary schools. *The Clearing House, 79*(1), 7-11.

Oday, K. A. (1984). *The relationship between principal and teacher perceptions of principal instructional management behavior and student achievement.* (Doctor of Education), Northern Illinois University, Dekalb, IL.

OECD. (2009). Creating Effective Teaching and Learning Environments: First Results from TALIS: Teaching and Learning International Survey.

Oesterle, M. (2008). *Student Perceptions of Effective Teaching.* (Doctor of Education), ARIZONA STATE UNIVERSITY.

Ostrander, L. P. (1995). *Multiple judges of teacher effectiveness: Comparing teacher self-assessments with the perceptions of principals, students, and parents.* (Ed.D.), University of Virginia, Charlottesville, VA.

Papanastasiou, E. (1999). *Teacher evaluation*. Michigan State University.

Perie, M., & Baker, D. (1997). *Job satisfaction among America's teachers: Effects of workplace conditions, background characteristics and teacher compensation.* Washington, DC: National Center for Educational Statistics.

Peske, G. H. & Haycock, K. (2006). Teaching Inequality: How Poor and Minority Students Are Shortchanged on Teacher Quality. *Education Trust*. June 2006, Washington, DC.

Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices.* Thousand Oaks, CA.: Corwin Press.

Peterson, K. D., Wahlquist, C., & Bone, K. (2000). Student Surveys for School Teacher Evaluation. *Journal of Personnel Evaluation in Education, 14*(2), 135-153.

Pianta, R., Mashburn, A., Downer, J., Hamre, B., & Justice, L. (2008). Effects of Web-Mediated Professional Development Resources on Teacher-Child Interactions in Pre-Kindergarten Classrooms. *Early Childhood Research Quarterly*, 23(4), 431-451.

Picus, L. O., Marion, S. F., Calvo, N., & Glenn, W. J. (2005). Understanding the relationship between student achievement and the quality of educational facilities evidence from Wyoming. *Peabody Journal of Education, 80*(1), 71-95.

Polikoff, S. M. (2015). The Stability of Observational and Student Survey Measures of Teaching Effectiveness. *American Journal of Education*, 121(2), 183-212.

Race to the Top Progress, sub-criterion, (2013). Part B Narrative, North Carolina.

Raudenbush, S. (2013). What do we know about using value-added to compare teachers who work in different schools? Carnegie Foundation for the Advancement of Teaching. Stanford, California.

Reichardt, R., Snow, R., Schlang, J., & Hupfeld, K. (2008). Overwhelmed and Out: A Research and Policy Report from Principals, District Policy, and Teacher Retention. Connecticut Center for School Change. http://www.nctq.org/nctq/research/1220022778926.pdf

Rice, J. K. (2003). *Teacher quality: understanding the effectiveness of teacher attributes*. Washington, D.C.: Economic Policy Institute.

Rice, J. K. (2009). Investing in Human Capital through Teacher Professional Development. *Creating a New Teaching Profession*. Washington, DC: The Urban Institute.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rockoff, J. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review: Papers and Proceedings of the One Hundred Sixteenth Annual Meeting of the American Economic Association, 94*(2), 247-252.

Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.

Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *Quarterly Journal of Economics, 125*(1), 175-214.

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

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*: John Wiley & Sons.

Sanders, W., & Horn, P. (1995). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. *Education Policy Analysis Archives, 3*(6).

Sanders, W., & Horn, S. P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247-256.

Sanders, W., & Rivers, J. (1996). Cumulative and residual effects of teachers on future student academic achievement. Knoxville, TN.: University of Tennessee Value-added Assessment Center.

Sanders, W., & Rivers, J. (1996). Cumulative and residual effects of teachers on future student academic achievement. Knoxville, TN.: University of Tennessee Value-Added Assessment Center.

Sanders, W., Wright, W., & Horn, S. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 4*(1), 3-7.

Sass, T., Hannaway, J., Xu, Z., Figlio, D., & Feng, L. (2012). Value Added of Teachers in High-Poverty Schools and Lower Poverty Schools. *Journal of Urban Economics, 72*(2-3), 104-122.

Scafidi, B., Sjoquist, D. L., & Stinebrickner, T. R. (2007). Race, poverty, and teacher mobility. *Economics of Education Review, 26*(2), 145-159.

Schneider, M. (2002). Do school facilities affect academic outcomes? National Clearinghouse for Educational Facilities, Washington, D.C.

Schneider, M. (2003). Linking School Facility Conditions to Teacher Satisfaction and Success. Washington, DC: National Clearinghouse for Educational Facilities.

Schumacker, R. E., & Lomax, R. G. (2010). *A Beginner's Guide to Structural equation Modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schochet, P. Z., & Chiang, H. S. (2010). *Error Rates in Measuring Teacher and School Performance Based on Student Test Score Gains* (NCEE 2010-4004). Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/pubs/20104004/pdf/20104004.pdf

Short, G. L. (1987). *The effects of the comprehensive behavior management system on student achievement, instructional climate, and the management of student behavior in selected junior high and middle schools in Texas.* (PhD), Texas A&M University.

Simon, N. S., & Johnson, S. M. (2013). *Teacher Turnover in High-Poverty Schools: What We Know and Can Do?* Project on the Next Generation of Teachers, Harvard Graduate School of Education.

Soebari, T., & Aldridge, J. (2015). Using student perceptions of the learning environment to evaluate the effectiveness of a teacher professional development programme. *Learning Environments Research*, 1-16. doi: 10.1007/s10984-015-9175-4

Sook-Jeong, L. (2007). The relations between the student-teacher trust relationship and school success in the case of Korean middle schools. *Educational Studies, 33*(2), 209-216.

Stachler, W. M., Young, R. B., & Borr, M. (2013). Sustainability of professional development to enhance student achievement: a shift in the professional development paradigm. *Journal of Agricultural Education, 54*(4), 13.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* Mahwah, NJ: Lawrence Erlbaum Associates.

Stronge, J. H. (2002). *Qualities of effective teachers*. Alexandria, VA: Association for Supervision and Curriculum Development.

Swank, P., Taylor, R., Brady, R., & Frieberg, T. (1989). Sensitivity of classroom observation systems: Measuring teacher effectiveness. *Journal of Experimental Education, 57*(2), 171-186.

Swanlund, A. (2011). *Identifying working conditions that enhance teacher effectiveness: The psychometric evaluation of the Teacher Working Conditions Survey.* Chicago. IL: American Institutes for Research.

Sykes, G. (2008). An inquiry concerning Teacher Working Conditions. Paper prepared for the Center for Teaching Quality, Hillsborough, NC.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics*. Boston: Allyn & Bacon.

Tennessee Department of Education. (2007). Tennessee's most effective teachers: Are they assigned to the schools that need them most? Research Brief. .

Tennessee State Board of Education. (2011). *Report card on the effectiveness of teacher training programs.* Report prepared in conjunction with the TDOE and the Tennessee Higher Education Commission.

Thomas, I. A., Green, R. L. (2015). Using Instructional Strategies to Enhance Student Achievement. *National Forum of teacher Education Journal*, 25(3), 1-18.

Tomaken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology, 1*, 2.1-2.35.

Tyler, J. H. (2011). *If you build it will they come? Teacher use of student performance data on a web-based tool.* NBER Working Paper.

Value-Added Research Center (2010). NYC Teacher Data Initiative: Technical Report on the NYC Value-Added Model. *Wisconsin Center for Education Research and New York City Department of Education*.

Vandenberg, R. J., & Nelson, J. B. (1999). Disaggregating the motives underlying turnover intentions: When do intentions predict turnover behavior? *Human Relations, 52*(10), 1313-1336.

Vogt, W. (1984). Developing a teacher evaluation system. *Spectrum, 2*(1), 41-46.

Ware, H. & Kitsantas, A. (2007). Teacher and collective efficacy beliefs as predictors of professional commitment. *The Journal of Educational Research*, 100(5), 308-321.

Watkins, P. (2005). The principal's role in attracting, retaining, and developing new teachers: Three strategies for collaboration and support. *The Clearing House*, 79(2), 83-87.

Wei, R., Darling-Hammond, L., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the U.S. and abroad*. Dallas, TX: National Staff Development Council.

Weiss, I. R., & Boyd, S. E. (1990). *Where are they now?: A follow-up study of the 1985-86 science and mathematics teaching force.* Chapel Hill, NC: Horizon Research, Inc.

Weiss, I. R., Pasley, J. D., Shimkus, E. S., & Smith, P. S. (2004). Looking inside the classroom: Science teaching in the United States. *Science Educator, 13*, 1-65.

Wenglinsky, H. (2000). *How teaching matters: Bringing the classroom back into discussions of teacher quality.* Princeton, NJ.: The Milken Family Foundation and Educational Testing Service.

White, M., & Rowan, B. (2014). *User Guide to measures of Effective Teaching Longitudinal Database*. The University of Michigan.

Wilkerson, D. J., Manatt, R. P., Rogers, M. A., & R., M. (2000). Validation of Student, Principal, and Self-Ratings in 360° Feedback (registered) for Teacher Evaluation. *Journal of Personnel Evaluation in Education, 14*(2), 179-192.

Wilson, S. M., & Floden, R. (2003). *Creating effective teachers: Concise answers for hard questions. An addendum to the report "Teacher preparation research: current knowledge, gaps, and recommendations."*. Washington, DC: AACTE Publications.

Yoon, K. S., Duncan, T., Lee, S. Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement.* Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.

Yuan, Y. C. (2010). *Multiple imputation for missing data: Concepts and new development*. Rockville, MD: SAS Institute Inc.