Running head: Measuring Teacher Quality

Measuring Teacher Quality:
Continuing the Search for Policy-Relevant Predictors of Student Achievement

Robert C. Knoeppel, Ph.D.
Associate Professor & Chair
Faculty of Leadership, Counselor Education, Human & Organizational Development
332 Tillman Hall
Box 340707
Clemson, SC 29631-0707
864-656-1882
864-656-1322 FAX
rck@clemson.edu

Joyce P. Logan, Associate Professor
Department of Educational Leadership Studies
College of Education
University of Kentucky
111 Dickey Hall
Lexington, KY 40509
859-257-7845
859-257-1015 FAX
robert.knoeppel@uky.edu

Clare M. Keiser, Ed.D.
Principal, Hollymead Elementary School
Charlottesville, VA 22903
ckeiser@k12albemarle.org

Paper originally presented at the annual conference of the
University Council for Educational Administration,
Nashville, TN
November, 2005

Measuring Teacher Quality:

Continuing the Search for Policy-Relevant Predictors of Student Achievement

The attraction and retention of high quality teachers are believed to be the most important policy alternatives by which schools can improve student achievement. Indeed, a growing number of research studies have identified the teacher as the single most important school-related input to improve student achievement (Cawelti, 1999; Jordan, Mendro, & Weerasinghe, 1997; Kaplan & Owings, 2002; Sanders & Rivers, 1996; Wright, Horn, & Sanders 1997; Darling-Hammond, 1997). While there has been general agreement that specific teachers have a strong, positive impact on student achievement, there is less agreement about effects of specific teacher attributes (Loeb, 2001).

Current literature on teacher quality can be grouped in three categories: preservice qualities of teachers, teacher practice, and teacher impact on student achievement. The predominant methodology used in these studies is the education production function (King Rice, 2003). These studies drive policy decisions for resource allocation as an attempt to maximize student performance. However, they have been limited in terms of measurable variables for teacher quality to indicators such as teacher certification, performance on certification exams, years of experience, relationship of teaching assignment to college major (Rebell & Wardenski, 2004) and student-teacher ratio (Krueger, 1999; Rivkin, Hanushek, & Kain, 2005). The least studied measure of teacher quality is teacher practice. Teaching practice is rarely examined as part of teacher quality research (Smith, Desimone, & Ueno, 2005). “A critical part of determining the

appropriateness of different observable indicators of teacher quality is examining the relationship of the indicators with different types of instruction to allow determination of the extent to which the indicators are proxies for 'better' instruction" (p. 76).

Value-added research is one approach that has been used to measure school effects on student learning (e.g. Sanders & Rivers, 1996). Loeb (2001) notes that this method defines teacher quality in a way that is of most interest to the public: student achievement gains. Potential problems with this method are nonrandom teacher assignment that may confound results (Loeb, 2001) and the complexity of the research design. Nonrandom teacher assignment may cause the impact of the teacher to be underestimated because some teachers may be more likely to teach students who have difficulty learning. A report by RAND researchers extensively reviewed the value-added model (Southeast Center for Teaching Quality, 2004) and cautioned against its use to rate individual teachers or as the sole basis for high-stakes school decisions. They also caution that current studies using this model do not always account for bias from school context, effect of previous teachers, or missing data.

Although value-added research adds to the knowledge about school effects on learning and, through statistical techniques, attempts to remove factors outside of school that relate to student achievement, this method places a heavy burden on identifying outcome-based measures of student achievement that discern the impact of teachers. Curriculum-based tests may not be a valid measure of teacher impact (Loeb, 2001). What is needed is an understanding of specific teacher and teaching characteristics

associated with student learning that during the course of instruction influences achievement levels of students (Odden, Borman, & Fermanich, 2004).

The purpose of this study is to extend the knowledge base about policy-relevant variables for teacher quality and their relationship to student achievement. Policy-relevant variables are those attributes of teaching quality that are measurable and can be advanced by policies to attract and retain high quality teachers in the nation's classrooms. The research described in this article examines teacher variables currently supported in the teacher-quality literature as being associated with student achievement and assesses a new variable: teacher certification by the National Board for Professional Teaching Standards (NBPTS). Goldhaber (2006) introduces this variable as an indicator for teacher quality based on empirical research that has documented student achievement gains made by students taught by National Board Certified Teachers (NBCTs). This variable relates to teacher performance and classroom practices and may help explain more of the school effect variance in student achievement.

Two research questions were posed in the study. First, are measures of teacher quality predictors of student achievement as measured by the Commonwealth Accountability Testing System (CATS) Index Score? Second, does the inclusion of a new measure of teacher quality that may act as a proxy for teaching practice, add to previous research that has found measures of teacher quality to be significant predictors of student achievement? National Board certification is a measurable variable that from its beginning in 1987 aimed toward advancing educational reforms for improving student achievement (Bond, Smith, Baker, & Hattie, 2000).

In 1990 the Commonwealth of Kentucky initiated one of the most sweeping educational reform measures in the country with the sole intent to improve education performance and results by addressing equity and equality issues (Foster, 1999). The Commonwealth Accountability Testing System (CATS) measures student performance by collecting academic and non-academic data and weighting those data to calculate an Accountability Index for each school. Using Kentucky as the context for studying teacher-quality product variables is particularly appropriate because of the wealth of data available over more than a decade. Reform and change in Kentucky schools are well established. Also, Kentucky implemented a teacher quality initiative for differentiated pay in 2003. National attention is focused on improving teacher quality and assigning the best teachers for at-risk children (Berry, 2004; Goldhaber, 2006; Humphrey, Koppich, & Hough, 2004). As an early implementer of systemic school reform, Kentucky is an appropriate site for study of teacher quality and student learning.

Rationale and Theoretical Framework

Education reform increased and reallocated resources to improve schools. Standards-based accountability systems were put into place to measure school results and ensure that dollars spent produced effects where it counts most—student learning. However, research efforts have begun to link the accountability of policies and dollars spent with student results (Downes & Figlio, 1999). In Kentucky, the link between accountability and student results was a highlight of the *Rose* decision that initiated education reform. After over a decade of experience with school reform implementation, evaluating such linkages increases in importance and interest to the public. According to

Stiefel, Schwartz, Rubenstein, and Zabel (2005), school performance and efficiency measures are at the center of educational policy debate and research and that efficiency measurement assume connections between inputs and outputs.

Relying largely on input measures for teachers and output (achievement) measures for students has advantages in terms of concreteness. Data on input are more accessible, reliable, and amenable to quantitative assessment than are data on the teaching practice that takes place between input and output—the quality of applying instructional skills and knowledge in a classroom context to advance learning for all students. Considerable research can be found on teacher quality and also on specific teaching strategies and practices that link to learning, particularly for specific subjects or groups of students. However, much remains to be done in teacher quality research to identify measurable teaching practice variables and their interaction with teacher, student, and school variables that can serve as policy-relevant predictors of student achievement.

National Board certification for teachers and its application process to measure teaching skills has potential as a measurable variable that could serve both as an input (the number of teachers with this certification) and a process (teaching practice) variable. The National Board for Professional Teaching Standards (NBPTS) states on its website that National Board certification "seeks to identify and recognize teachers who effectively enhance student learning" (NBPTS, n.d.b, "About NBPTS," Policy Position section, para.1). Although not all research on NBPTS certification has shown this certification related to student learning (McColskey, 2006; Sanders, Ashton, & Wright, 2005; Stone, 2004), studies are emerging that suggest a degree of NBPTS success in

identifying high quality teachers associated in with student achievement, at least in some subjects and grade levels (Goldhaber, 2006; Humphrey, Koppich, & Hough, 2004; O'Sullivan et al., 2005; Vandervoort, 2004), and all 50 states offer bonuses and/or incentives for this national certification (NBPTS, 2004a).

Schwartz and Stiefel (2004) discuss the difficulty of measuring educational inputs and outputs, capturing contextual influences, compensating for data scarcity, and attributing causality. However, they see promise in improved measures of school efficiency and potential for identifying determinants of efficiency across schools and school districts.

Application of selected policy-relevant variables to Kentucky's accountability system explores not only how previously identified measurable teacher quality variables apply to Kentucky schools, teachers, and student outcomes but also envisions advancing teacher quality research beyond teacher demographics such as pre-service preparation and experience brought to the classroom. Figure 1 illustrates a framework for extended research on teacher quality that represents the complex nature of teaching.

[Insert Figure 1 about here]

Teacher quality is the ability to interact with students and advance what students know and are able to do. However, the value added to learning by teachers depends on what they know, what they do, and how they interact within the context of their work. Learning-related components of teacher quality may be grouped as teacher demographics, teaching practice, and organizational context. These components interrelate to achieve school outcomes.

No Child Left Behind legislation defines the minimum teacher quality requirements as state certification, a bachelor's degree, and state-defined demonstrated competence in core subjects taught (U.S. Department of Education, n.d.). Demonstrated competence implies teaching practice but has no common definition for states. Teacher quality must be more broadly defined than degrees, state certification, and teaching experience. These variables help define teacher quality but fall short of addressing the complexity of teaching and learning.

Reeves (2006) proposes that effective teaching practices are widely known but not widely practiced; he also describes education as complex and multivariate with multiple measures of student achievement. Teacher demographics have been studied, although with mixed result, as predictors of student achievement. Policy-relevant variables must be measurable and achievable by policy directive. Thus, state certification requirements and salary structures that address experience and degrees are widely used as an effort by states to affect the quality of teachers. However, what teachers know about teaching and learning must be evidenced also by teaching practice. Teaching practice is what teachers do in the classroom to advance student achievement. As defined for this paper, high quality teaching practice means effective application of subject content and pedagogical knowledge and skills consistent with student needs and research-based practices that advance student achievement within a specific classroom and school context.

Standards-based teacher evaluation is a policy-directed attempt to improve and measure specific teaching practice. Although quantifiable, teacher evaluation retains

evaluator subjectivity, making it more amenable as one of multiple measures of teaching practice. Also, questions arise about the effectiveness of teacher evaluation in achieving quality teaching practice. In addition, volumes of research identify specific teaching practices that relate to student learning (Marzano, Pickering, & Pollock, 2001); for example, knowledge application through problems, projects, and simulations (Wolfe, 2001); authentic assessments (Wiggins, 1993); and classroom behavior (Tobin & Sugai, 1999). However, the combination and application of best practices within a classroom for student achievement are not as easily identified. Cooley and Leinhart (1978) and Brophy and Good (1986) studied effective teachers as measured against standardized test gains and found such things as planning, making goals clear to students, and using class time well with coherent instructional strategies; these studies preceded today's knowledge of teacher effect sizes related to learning. Cohen, Raudenbush, and Ball (2003) recognize the interconnectedness of teachers and students and propose that teachers adjust instruction "to their view of students' capabilities," and students "calibrate their use of resources to their estimates of teachers' and parents' expectations" (p. 133). This interaction affects teaching practice, as do interactions with school leadership and other school resources.

Teacher certification by the National Board for Professional Teaching Standards, in one sense, is a demographic variable because it requires three years of experience and state certification to apply. But the process of identifying teachers for this national recognition of teaching quality places emphasis not only on what they know but also what they do in their classroom, as identified through videotapes of teaching practice in

the context of school and classroom plus a portfolio that describes teaching practice and reflects on it related to standards. Although research studies show mixed results about the relationship of National Board certification to student achievement, research lends support for NBPTS accuracy in identifying quality teachers and a relationship to student achievement in specific subjects or depth of learning (Bond et al., Vandervoort, Amrein-Beardsley, & Berliner, 2004). Thus, NBPTS seems to merit further study as one proxy for teaching practice. Whether or not a teacher has NBPTS certification is measurable, and increasing the number of teachers with this specialized certification is achievable by policy directive.

Student achievement as defined in Kentucky's accountability system uses multiple measures of student success. Test scores are an essential component (both national standardized tests and tests of Kentucky Core Content), but student success measures are also high school completion and transition success of students to further education and/or employment, as well student behaviors as evidenced by school attendance and school completion. More teacher quality research has been conducted with elementary school students than high school; but when extended to middle and high school students, success can be evidenced not only by test scores but also how they apply their learning after graduation.

Related Literature

The No Child Left Behind Act of 2001 and the Teaching Commission report (2004) increased accountability for student performance and reinforced the effort to place the best teachers possible in all classrooms. The recent focus on recruiting and retaining

high quality teachers stemmed, at least in part, from a number of research studies that reported the teacher as the single most important school effect on student learning (Cawelti, 1999; Johnston, 1999; Jordan, Mendro, & Weerasinghe, 1997; Sanders & Rivers, 1996; Wright, Horn, & Sanders, 1997). The quality of teaching matters (Rivkin, Hanushek, & Kain, 2005; Sanders & Rivers, 1996; Stronge & Tucker, 2000).

Students taught by good teachers progress academically at higher rates than do students in classrooms with poor teachers (Sanders, 2000; Sanders & Horn, 1998; Sanders, Wright, & Ross, 1999; Topping & Sanders, 2000). Furthermore, teacher effects on student learning are additive and cumulative over grade levels (Sanders, 2000).

The distribution of high quality teachers across schools and districts varies. Teachers with less experience, education, and lower test performance are most likely to work in high minority, high poverty, and low-performing schools (Knoeppel, 2007; Wyckoff, 2003). Additionally, teacher quality has a greater impact on poor students than on high income students (Coleman, 1990; Goldhaber & Anthony, 2004). Therefore, identification of teacher quality and quality teaching practice is important not only for teacher recruitment but also to assure equitable distribution of excellent teachers for at-risk children.

Teacher quality is most often defined by presage variables when investigating effects on student achievement (Rowan, Correnti, & Miller, 2002) and has been examined by the use of production function research. King Rice (2003) summarized empirical research evidence related to variables representing teacher characteristics that are measurable and important to teacher quality, including teaching experience, teacher

preparation programs and degrees, and teacher coursework. These areas relate directly to the selection of variables used in the Kentucky research reported in this article.

Ferguson and Ladd (1996), Greenwald, Hedges, and Laine (1996), and Murnane and Phillips (1981) studied effects of teaching experience as a measurable variable correlated with student achievement. Based on a review of these studies, King Rice (2003) concluded that teaching experience appears to have a relationship to student achievement. Archibald (2006) concurs, but she notes that findings with regard to this experience are mixed. Positive relationships have been found between teacher experience and student achievement (Hanushek, 1992, 1997; Rowan, Correnti, & Miller, 2002), but these benefits are found after the first few years teaching (Hanushek, Kain, & Rivkin, 1998; Murnane, 1983; Rockoff, 2004). Teachers with less than three years of experience have been found to be less effective than their more experienced counterparts (Darling-Hammond, 1999). Teachers become more effective during their initial years of experience, and benefits tend to level off after about five years of experience. However, the relationship of teacher experience to student achievement is not linear, Veteran teachers do not always continue to grow; but in settings that emphasize continuous learning and collaboration, veteran teachers are more likely to continue to improve performance (Rosenholz, 1986).

Research on the value of a teacher's advanced degree is also mixed. Almost all teachers now have an undergraduate degree, and more teachers' highest degree is a master's rather than a bachelor's degree (Wyckoff, 2003). Some studies show that additional teacher education has a positive correlation with student achievement; others

studies do not. Goldhaber and Brewer (1997) found that from the eighth to the tenth grade, teachers' advanced degrees are not generally associated with increased student achievement. The same study found that teachers with advanced degrees in mathematics and science appear to have a positive influence on student achievement. Results from all studies collectively suggest that a teacher's advanced degree has a positive correlation with student achievement only when that degree is in subjects taught (Goldhaber & Anthony, 2004). Although overall the evidence is weak for the benefit of advanced degrees for improving student achievement, positive benefits do seem to occur at some levels and for certain subject areas.

Studies examining the relationship between teacher coursework and student achievement have produced mixed results, although subject content knowledge seems to matter (Goldhaber & Brewer, 1997; Hawk, Coble, & Swanson, 1985; Monk & King, 1994). However, after a threshold of competency is attained, pedagogical teacher training may be more important to student success than content knowledge (Laczko-Kerr & Berliner, 2002). Teacher coursework, both in pedagogy (Adams & Krockover, 1997; Ferguson & Womack, 1993; Monk & King, 1994) and the subject taught (Monk & King, 1994), has been found to relate to positive student results. Course content in the subject taught appears most important at the high school level.

Some studies have employed the variable student-teacher ratio to study how class size affects education quality, although acknowledging that student-teacher ratio is not the same as the number of students taught by a teacher because the ratio is calculated using all certified staff in a school. However, use of the ratio rather than actual number of

students assigned to each teacher is a limitation of these studies. The studies showed significant effects of class size for both mathematics and reading achievement gains, but the effect declined as students progressed through school (Rivkin, Hanushek, & Kain, 2005). Strong evidence exists for the benefits of small class size for kindergarten and through grades 1 to 3 (Word, Johnson, Bain, et al., 1990). These students in small classes demonstrated substantially higher levels of reading achievement, and the gains were greater the longer they were exposed to small classes.

A longitudinal study in Tennessee followed kindergarten students assigned to small classes from K-3. Results showed that the students assigned to small classes maintained an academic advantage at least to the eighth grade (Nye, Hedges, & Konstantopoulos, 2000). This Tennessee study, Project STAR (Student/Teacher Achievement Ratio), was a controlled scientific experiment that followed kindergarten students randomly assigned to a small class (13-17), a regular class (22-26), and a regular class with full-time teacher aide (Finn & Achilles, 1999; "Reducing Class Size," 1999). At the end of kindergarten, the small-class students were .7 to .9 months ahead of the other two groups in all subjects. After first grade, small-class students were about 2 months ahead. At the end of the fifth grade (although students returned to full-size classes in Grade 4) the small-class students were 5 months ahead of regular-class students in all subjects (Finn & Achilles, 1999, p. 1-101). In a Grade 4 follow-up study, specific learning behaviors were assessed that measured effort in learning, initiative in the classroom, and nonparticipatory behavior, students who had been in the small-class group K-3 had a carryover effect in that a year after being returned to full-size classes, those

who had been in the small-class group rated superior on all of the behavior scales. Results from this rigorous study and studies in Wisconsin and North Carolina showed positive results in student learning with small-size classes ("Reducing Class Size," 1999).

Significant effects of class-size reduction appear when the number of students is from 15 to 20 students and continue increasing as class size approaches a one-to-one tutorial. Teachers report that the classroom environment is better when reducing class size. Students receive more attention and participate more, and the teachers know the students better than in a regular class and develop better relationships with them. Although not a magic number, research makes it fairly clear that class size must get below 20 to make a difference. The greatest benefit of small class size appears with groups of students who are disadvantaged (Biddle & Berliner, n.d.). Research for grades 4 through 12 is less certain about its effects than is true for K-3.

Critics have noted that many "class-size" studies examine student-teacher ratio, rather than actual class sizes, which ignores how students are assigned to classrooms because ratio is based on all certified adults in the school rather than actual student class assignments. However, the STAR project and other experimental research designs used random class assignments. Even with randomly assigned classes, student in- and out-migration likely did not hold per class numbers constant throughout three years, but this research is considered the "gold standard" of the small class-size research.

National Board Certification

Founded in 1987, the National Board for Professional Teaching Standards had a three-fold mission: (a) advance the quality of teaching and learning by maintaining high

and rigorous standards for what accomplished teachers should know and be able to do, (b) provide a national voluntary system certifying teachers who meet these standards, and (c) advocate related education reforms to integrate National Board certification in American education and to capitalize on the expertise of teachers certified by the National Board (NBPTS, n.d.a.). Believing that the most effective way to increase student learning was to improve teacher practice, the National Board sought to focus education reform initiatives on the teacher (Bond, Smith, Baker, & Hattie, 2000). It was intended that this certification would create an identifiable group of teachers who would then motivate the teaching profession by giving back to the profession and by serving as teacher leaders and mentors in their respective schools (NBPTS, n.d.a.). The National Board believed that the adoption of standards that represented accomplished teaching as well as the creation of a reliable and valid system of assessment would promote realization of their vision.

Five core propositions were articulated by the National Board. They included the following: teachers are committed to students and their learning, teachers know the subjects they teach and how to teach these subjects to students, teachers are responsible for managing and monitoring student learning, teachers think systematically about their practice and learn from experience, and teachers are members of learning communities (Harman, 2001). Building on these five core propositions, the National Board designed specific content standards for each field of certification. Twenty-four certificate fields were available as of March 2005. Applicants for national certification must meet course level and subject area standards that were developed by committees of teachers and other

experts to demonstrate what the most accomplished teaching professionals are able to do in the classroom. This certification is a voluntary process but is not available to teachers until they have earned a baccalaureate degree, have a minimum of three years teaching experience and a valid state teaching license for those years taught (NBPTS, 2004, p. 11). Where a license is not required, their teaching must have been in schools recognized and approved by the state where they taught.

Candidates for National Board Certification are assessed both by a portfolio documenting how their teaching practice meets NBPTS standards (American Federation of Teachers & National Education Association, 2005) and testing in an assessment center. The assessment center test requires successful completion of responses to computer-delivered prompts. These prompts are designed to elicit knowledge and skills related to the teacher's content field and level of certification (NBPTS, 2004, p. 8). The portfolio, prepared by the NBCT candidate over time, has four entries that focus on classroom practice and usually include two videotapes of teaching along with student work samples. A reflective analysis by the teacher accompanies each entry explaining how his or her teaching practice meets NBPTS standards.

Goldhaber (2006) introduces this variable as an indicator for teacher quality based on empirical research that has documented student achievement gains made by students taught by National Board Certified Teachers (NBCT). This variable relates to teacher performance and classroom practices and may help explain more of the school effect variance in student achievement. Because the certification process has been identified as a means to identify accomplished practice, an emerging body of literature has examined

the validity of this designation (McColskey, Stronge, et al, 2006). These studies have explored the relationship between Board certification and student outcomes as measured on state tests and teacher practice.

The strongest support to date for the effect on student learning comes from a large-scale study in North Carolina (Goldhaber & Anthony, 2004). This study found that in "both reading and mathematics, NBCTs were more effective than non-certified applicants" and concluded that "NBPTS is in fact identifying the more effective of those teachers that they actually evaluate" (p. 17). This study suggests that NBPTS certification conveys teacher quality information for performance beyond that of teacher licensure.

A large-scale study in Arizona that analyzed four years of reading, mathematics, and language scores and tested for statistical difference supports Goldhaber and Anthony's findings (Vandervoort, Amrein-Beardsley, & Berliner, 2004). Results from this research found National Board Certified Teachers' (NBCTs') students outperformed non-NBCTs' students on 72.9% of the measures over a four-year period. Some students of NBCTs had over one month advantage in achievement compared to non-NBCTs' students. Effect sizes were greater in mathematics and reading than in language.

An early study of NBPTS certification that considered teacher practice compared teachers who successfully attained national certification with teachers who unsuccessfully attempted this certification (Bond, Smith, Baker, and Hattie, 2000). This validity study of the NBPTS process used onsite visits to classrooms, interviews with teachers and students, and samples of student work. In this sample, National Board Certified Teachers demonstrated expert teaching to a greater degree than did the non-

certified teachers. They also found that at the high school and middle school levels student understanding of unit objects was richer, more elaborated, and more interconnected than teaching outcomes for non-certified teachers, although writing samples showed only marginal differences.

Elfers and Plecki (2006) found that NBCTs report greater understanding of curriculum, instruction, and assessment after certification than before. Some of the studies reported that these teachers demonstrated greater ability to manage diverse learning needs in their classroom after successfully completing the certification process (Elfers & Plecki, 2006; Vandervoort, Amrein-Beardsley, & Berliner, 2004). This may have implications for teacher quality equitable distribution-- particularly because the best teachers can make the most difference for at-risk students, yet usually choose to teach in high performance schools.

Aside from the mixed findings in these studies, the research has been criticized on methodological and statistical issues, most notably small samples that lack statistical power; large samples that result in statistical significance but not meaningful differences; designs that fail to consider student attributes and the correlation between these attributes and teacher assignment; and lastly, inaccurate links between student data and teacher assignment (McColskey, Strong, et al, 2006). The fact that most NBCTs are not placed in hard-to-staff schools (Goldhaber, 2006; Humphrey, Koppich, & Hough, 2004) also may skew study results.

This study attempts to contribute to the extant literature by including student demographics in the research design. Further, the findings include a discussion of effect size and power.

Method and Results

This study sought to extend the literature base on the relationship between measures of teacher quality and student achievement through the inclusion of a new variable: National Board Certified Teachers. Two research questions were posed in the study. First, are measures of teacher quality predictors of student achievement as measured by the Commonwealth Accountability Testing System (CATS) Index Score? Second, does the inclusion of a new measure of teacher quality that may act as a proxy for teacher practice, add to previous research that has found measures of teacher quality to be significant predictors of student achievement? Previous studies examining the relationship between measures of teacher quality and student achievement have focused on measurable, policy relevant variables. These variables have not reflected measurable teacher behaviors in the context of teaching and learning.

Education production functions have been the predominant methodology utilized to discern if a relationship exists between resources and measures of student achievement. The production function is a statistical technique that describes the maximum level of outcome possible from different combinations of inputs. An example of a production function that utilizes a statistical technique to analyze the relationship between school resources and student learning is multiple regression analysis. In multiple regression, a linear combination of independent variables is formed by

differentially weighting each independent variable so that the correlation with the dependent variable is maximized (Stevens, 1992). Sheskin (2000) states that a goal of multiple regression is to identify a limited number of predictor variables that optimize one's ability to predict scores on the dependent variable. Through regression, the researcher is best able to make a prediction of a given phenomena (Pedhazur, 1982).

This analysis includes two distinct purposes, *correlation* and *regression*, even though the terms are used interchangeably. First, regression analysis is a technique to find the relationship between one dependent variable and two or more independent variables, which is multiple correlation (Pedhazur, 1997; Tabachnick & Fidell, 2007). A second pupose is to predict future outcomes based upon analyzing an outcome measure from several independent variables. Both purposes can be utilized in interpreting the outcomes when multiple regression is used to analyze production function data (Tabachnick & Fidell, 2007: Pedhazur, 1997; Stevens, 1996). One use of multiple regression in education is to explain student learning based upon inputs found in school settings (Pedhazur, 1997). Another important output from multiple regression analysis is the correlation between the independent variables and the dependent variable. This relationship is known as the squared multiple correlation coefficient (R^2) and indicates the amount of variance in the dependent measure accounted for by the independent variables. Typically, in social science research, the amount of variance explained is about 1% (Cohen, 1989).

The existence of a production function infers that there is something systematic about the transformation of inputs into outcomes (Steifel, Schwartz, Rubenstein, & Zabel,

2005; Schwartz & Zabel, 2005; Monk, 1989). Studies employing a production function have included inputs such as resources, organizational characteristics, and student attributes; outputs have included myriad measures of student achievement. These output measures may take the form of level scores, gain scores, or difference scores (Schwartz & Zabel, 2005). By including student demographics, the researcher is able to account for variance in student achievement that may be outside the control of the school, thus enhancing the ability to make inferences with regard to efficiency of resource allocation and effectiveness.

Sampling and Participants

This study employed a random sample of 339 schools in the Commonwealth of Kentucky and included school level measures of teacher quality and measures of student achievement. The sample size was guided by Cohen's (1989) work with regard to effect size, predicted relationship, and power analysis. Descriptive statistics appear in Table 1.

Study Design

To answer research questions posed in the study, a sequential multiple regression was performed between the dependent variable 2004 CATS scores and the independent variables percentage of students receiving special education services, percentage of students receiving services for free and reduced lunch, the percentage of students receiving services for limited English proficiency, assessed value per pupil, percentage of teachers with a major or minor in the content area taught, percentage of teachers with a master's degree, average years of experience, student-teacher ratio, and percentage of teachers who were a National Board Certified Teacher. Three separate groups of

variables were entered in the sequential multiple regression. The first group included measures of student demographics including the percentage of students receiving free and reduced lunch, the percentage of students receiving special education services, and the percentage of students receiving services for limited English proficiency. The second group included the variable assessed property value per pupil. This variable was utilized as a proxy for organizational characteristics. The variable assessed value per pupil refers to the ability of any given locality to provide financial support for education. This variable is of particular interest in Kentucky given the *Rose* decision that stipulated an equal, adequate education must be provided in the Commonwealth without regard to locality. For the purposes of this study, assessed value per pupil is also utilized as a measure of social capital of the school district. Independent variables in Models 1 and 2 were entered into the predictive model using standard linear regression. The third group included five measures of teacher quality: percentage of teachers with a major or minor in the content area taught, percentage of teachers with a master's degree, average years of experience, student-teacher ratio, and percentage of teachers who were a National Board certified teacher. These variables were entered into the predictive model using stepwise linear regression. In stepwise linear regression, independent variables that contribute significantly to the variance explained in the dependent variable are identified. A total of 6 models are identified in the sequential multiple regression. The first model includes the variables in group 1. Model 2 includes the variables in group 2. Models 3 through 6 identify measures of teacher quality that significantly contribute to and explain the variance in student achievement and are identified in order of importance.

For each independent variable in the third group, the term *teacher* refers to one who is teaching in either an individual or collaborative classroom. The percentage of classes taught by teachers with a major, minor, or the equivalent in the subject taught provides information on the extent of content preparation of teachers in each school in grades 6-12. Elementary school teachers with P-5 certification are counted in this percentage. The term *equivalent* is defined to include alternative certification, endorsements and passing the PRAXIS II content test, and middle school certification with areas of concentration. Percentage of teachers with a master's degree or greater reports the percentage of teachers in each school that has earned an advanced degree. The variable average years of experience reports the average number of years of teaching experience for teachers in each school. The variable National Board Certified Teachers refers to the percentage of teachers in each school who have earned National Board certification prior to the 2003-2004 school year. The student-teacher ratio variable refers to the average number of students in each class in the school for each teacher.

Multiple regression makes use of both correlation and regression. Correlational data are presented in Table 1. Pedhazur (1982) makes a compelling argument that the researcher needs to consider both the significance and the meaningfulness of the relationship described by a correlation coefficient. He reminds the researcher that "all that is meant by a statistically significant finding is that the probability of its occurrence is small, assuming that the null hypothesis is true" (Pedhazur, 1982 pp. 24-25) and that a statistically significant finding is not important if it is not meaningful. Emphasis should be placed on the magnitude of the relationships among variables and the magnitudes of

differences among means (Pedhazur, 1982). As such, this study included an analysis of the size of the relationships as well as an analysis of the outcome measures on the CATS index. According to these data, a statistically significant relationship at the $p \leq .01$ level was found to exist between the 2004 CATS Index and three measures of teacher quality: percentage of teachers with a major or minor in the content area taught, average years of experience, and percentage of teachers who were a National Board Certified Teacher. According to Cohen (1989), the effect size of these relationships would be considered to be a small effect, although all three relationships are approaching a medium effect. The variable percentage of teachers with a master's degree was found to have a statistically significant relationship at the $p \leq .05$ level with the 2004 CATS index. This relationship could also be classified as having a small effect size. Interestingly, the independent variable student-teacher ratio did not have a statistically significant relationship with the 2004 CATS index.

Results from the sequential multiple regression are presented in Table 2. Sequential multiple regression was the chosen method of analysis so that variance explained by student demographics could be separated from the variance explained by inputs to schooling in order that efficiency conclusions could be drawn. According to those data, student demographics significantly predict student achievement in Model 1, $R^2=.233$, $R^2_{adj}=.227$, $F(3, 335)=33.998$, $p \leq .000$. Model 1 accounted for 23.3% of the variance in student achievement as measured by the 2004 CATS index. Table 2 also displays the unstandardized regression coefficients (B), standardized regression coefficients (β), significance level of the regression coefficients, and tolerance for each

independent variable. These data enable the researcher to discern which independent variables were significant predictors of student achievement. Individually, the independent variables percentage of students receiving free and reduced lunch (t=-8.684, $p\leq$.000) significantly predicted student achievement in Model 1 as measured by the 2004 CATS index. Measures of tolerance calculated in the model indicate that redundancy is not a problem. Model 2 in the sequential multiple regression was not found to be a significant predictor of student achievement. Wealth of the locality did not significantly predict student achievement as measured by the 2004 CATS index.

Model 3 identifies the independent variable percentage of teachers with a major or minor in the content area as a significant predictor of student achievement and as the independent measure of teacher quality that explains the largest amount of variance in student achievement, R^2=.264, R^2_{adj}=.253, F(1, 333)=13.638, $p\leq$.000. Total variance explained by this variable was 3%. Model 4 identifies the input variable average years of experience as a significant predictor of student achievement and as a variable that explains a significant amount of variance in student achievement, R^2=.284, R^2_{adj}=.271, F(1, 332)=9.372, $p\leq$.002. The additional variance explained by the inclusion of this variable was 2%.

Model 5 identifies the variable student-teacher ratio as the next most important predictor of student achievement, R^2=.299, R^2_{adj}=.284, F(1, 331)=7.143, $p\leq$.008. It is interesting to note that the variable student-teacher ratio was not found to have a statistically significant relationship with the dependent measure 2004 CATS index; however, this variable has been found to explain 1.5% of the variance in student

achievement. Model 6 identifies the percentage of teachers holding National Board certification as the final significant predictor of student achievement, R^2=.313, R^2_{adj}=.297, $F(1, 330)=6.707$, $p \leq .010$. This measure of teacher quality was found to explain 1.4% of the variance in student achievement.

Conclusions and Implications for Future Research

Are measures of teacher quality significant predictors of student achievement? As educational leaders continue to grapple with state and national mandates to improve student achievement, this question becomes increasingly important. The informed, efficient allocation of scarce resources to support learning goals by educational leaders, especially for underrepresented populations, is the challenge at all levels of schooling. This study extends past research that has shown a relationship between student achievement and teacher quality. The review of the literature indicates that attempts to answer this question have been limited primarily to the use of measurable policy-relevant variables. Those studies have found small to moderate relationships between student achievement and measures of teacher quality. Hanushek, Kain, and Rivkin (1998) suggest that variations in teacher quality account for at least 7.5% of the variance in student achievement although they concede that the effect may be higher. Total variance in student achievement that was explained by measures of teacher quality in this study was 7.9%.

This study appears to substantiate to some degree previous research related to teacher and school effects and has broadened the base for comparison to student achievement by using Kentucky's CATS school indexes comprised of multiple measures

of tests and transition after-high school data. Not surprisingly, the largest effect size predictor of student achievement in this study was a major or minor in the content area. This is consistent with previous research on teacher quality in that knowledge of content has been identified often as a predictor, although it seems to be more important for certain subject areas or school levels. This study did not examine data by levels or subject area. The Master's degree showed a small effect, and previous research provides little support for advanced degrees, particularly if they are not in the content area.

Student-teacher ratio in this study did not show a significant effect size but did explain 1.5% of the variance in student achievement as measured by the CATS index. This mixed result is somewhat puzzling, but student-teacher ratio is not an accurate measure of class size. Kentucky Revised Statute 157.360 (Legislative Research Commission, 2006) sets a maximum number of pupils enrolled in a class to 24, K-3; 28, Grade 4; 29, Grades 5-6; and 31, Grades 7-12. Because previous research on class size suggests the threshold of benefit to be below 20 students and to be more clearly established for K-3, perhaps nearness of Kentucky's cap size for K-3 to the 20 threshold may help explain this finding. Kentucky's funding formula treats the class caps similar to calculation of student-teacher ratio by allocating funding for the number of total certified position based on student enrollment in grade-level categories compared to class cap sizes rather than to actual student assignment per teacher.

The number of National Board certified teachers in Kentucky represents a small percentage of all teachers, but the assumption (supported by the process for approval and a number of previous research studies) that NBCTs can be a proxy for teaching practice

merits further study. NBCTs is a measurable policy-relevant variable of particular interest not only because it moves the study of teacher quality into the classroom setting and interactions with students but also because incentives and/or bonuses for achieving this national recognition is a matter of policy in all 50 states (NBPTS, 2004a). However, as budgets have tightened and research has shown that fewer National Board certified teachers are found in schools with high minority population and low poverty schools, policy makers are taking another look at how to distribute these resources for national certification (Humphrey, 2004; "Quality Counts," 2003). California eliminated their $10,000 bonus for certification but maintained a $20,000 award for NBCTs working in low performing schools, and other states are beginning to take a look at their policies for distributing funds to place quality teachers in assignments where they are most needed. Therefore, now is an appropriate time to take a closer look at National Board certification as a predictor of student achievement and what this means for further research on the best use of this teacher quality resource.

In this Kentucky study, with the addition of National Board certification as a variable that served as a proxy for teacher practice and knowledge of pedagogy, the amount of variance in student achievement, after accounting for student demographics, was slightly higher than the findings of previous research.

Four measures of teacher quality included in this study (percentage of teachers with a major or minor in the subject area taught, years of teaching experience, percentage of teachers who hold National Board certification, and student-teacher ratio) were significant predictors of student performance on the CATS index score. Descriptive

statistics reveal that merely 1.80% of teachers in Kentucky's public schools have obtained National Board teaching certification. Given that the variable NBCT was a significant predictor of student achievement in the sequential multiple regression model, it would appear that a legally defined variable such as National Board certification has merit in terms of helping define teacher practice and the critical relationship to behavior and learning. This finding is somewhat mitigated by the fact that the relationship found between the percentage of teachers holding National Board certification and the 2004 CATS index is considered to be small. Although the relationship is statistically significant, the effect size of this relationship may make the use of this variable less practicable.

This finding suggests the need for more research in the area of teacher quality. Further research should examine the effect of NBCTs on student achievement and the effects of instructional pedagogy, as well as other school-classroom contextual evidences of the quality of teaching. What skills do successful teachers employ? How do their relationships and interactions with students impact achievement? In what ways do the current means by which we define quality teachers fail to explain what makes for good instruction? How are additive effects of multiple variables determined when teacher indicator variables interact or overlap?

Research-based decisions about teacher recruitment, selection, employment, and retention are critical for school leaders to locate, assign, and retain high quality teachers and teaching practices in every classroom. Research indicates that it is not only finding and hiring high quality teachers in schools but that keeping them relates to school

working conditions (Humphrey, 2004). The role of the school leader centers on teaching and learning and working with teachers and staff to create a school environment conducive to learning. “Effective teachers do not want to work with ineffective principals” (Southeast Center for Teacher Quality, 2003, p. 5). Therefore, leadership is a key factor in the teacher quality initiatives.

The authors of this article argue that teacher quality is more complex than the qualities that teachers bring to the classroom. Thus, further research should identify policy-relevant variables to serve as measures of relationships to student achievement. These measures should represent not only teacher demographics and characteristics but also teaching practice that interacts with school leadership, class size, availability of facilities and supplies, and other school and community effects on learning.

Because of the preponderance of research supporting the importance of teacher quality for student learning gains, such research is critical to increasing student achievement. Findings from teacher quality research can be used to guide best practice in teacher education and the preparation of instructional leaders. Moreover, studies examining the relationship between resources to schooling and measures of student achievement are of paramount importance to educational leaders as schools confront looming deadlines for proficiency for all students by 2014.

References

Adams, P.E., & Krockover, G.H. (1997, January). Concerns and perceptions of beginning secondary science and mathematics teachers. *Science Education, 81*(1), 29-51.

American Federation of Teachers & National Education Association. (2005). *A candidate's guide to National Board Certification*. Washington, DC: Authors.

Archibald, S. (2006). Narrowing in on educational resources that do affect student achievement. *Peabody Journal of Education, 81*(4), 23-42.

Berry, B. (2004, March). Recruiting and retaining "highly qualified teachers" for hard-to-staff schools. *NASSP Bulletin, 87*(638), 5-27.

Biddle, B.J., & Berliner, D.C. (n.d.). What research says about small classes and their effects. *WestEd Policy Perspectives*. Retrieved September 19, 2006, from www.wested.org/cs/we/print/docs/we/home.htm

Bond, L., Smith, T., Baker, W.K., & Hattie, J.A. (2000, September). *The certification system of the National board for Professional Teaching Standards: A construct and consequential validity study*. Washington, DC: U.S. Department of Education and the National Board for Professional Teaching Standards.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M.C. Witrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). MacMillian Publishing.

Cawelti, G. (1999). Improving achievement: Finding research-based practices and programs that boost student achievement. *The American School Board Journal, 186*(7), 34-37.

Cohen, J. (1989), *Statistical power analysis for the behavioral sciences* (2nd ed.), Hilldale, NJ: Lawrence Erlbaum Associates, Publishers.

Cohen, D.K., Raudenbush, S.W., & Ball, D.L. (2003). Resources, instruction, and research, *Educational Evaluation and Policy Analysis*, 25(2), pp. 119-142

Coleman, J.S. (1990). *Equality and achievement in education*. Boulder, CO: Westview Press.

Cooley, W., & Leinhardt, (1978). *Instructional dimensions study*. Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence.* University of Washington, Center for the Study of Teaching and Policy: A National Research Consortium

.

Darling-Hammond, L. (1997). Quality teaching: the critical key to learning. *Principal, 77*, 5-6.

Downes, T.A., & Figlio, D.N. (1999). Economic inequality and the provision of schooling. *Economic Policy Review, 5*(3) 99-111.

Elfers, A. M. & Plecki, M. L. (2006). *National board certified teachers in the Washington workforce: Views of teaching.* . Paper presented at the Annual Meeting of the American Education Finance Association, Denver, CO.

Ferguson, R., & Ladd, H. (1996). How and why money matters: An analysis of Alabama schools. In H.Ladd (Ed.). *Holding schools accountable*. Washington, DC: Brookings Institute Press.

Ferguson, R.F., & Womack, S.T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education, 44*(1), 55-63.

Foster, J. D. (1999). *Redesigning pubic education: The Kentucky experiences.* Lexington, KY: Diversified Services.

Goldhaber, D. (2006). National board teachers are more effective, but are they in the classrooms where they're needed the most? *Education Finance and Policy, 1*(3), 372-382.

Goldhaber, D., & Anthony, E. (2004). *Can teacher quality be effectively assessed?* Retrieved November 26, 2004, from http://www.crpe.org/workingpapers/pdf/NBPTSquality_report.pdf

Goldhaber, D. D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*, 129-145.

Goldhaber, D.D., & Brewer, D.M. (1997). Evaluating the effect of teacher degree level on educational performance. In Fowler, J.W. (Ed.). *Development of school finance*. Washington, DC: National Center for Education Statistics.

Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Education Research, 66*(3), 1141-1178.

Hanushek, E.A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy analysis, 19*(2), 141-164.

Hanushek, E.A. (1992). Assessing the effects of school resources on student performance: An update. *Journal of Political Economy, 100*(1), 84-117.

Hanushek, E.A., Kain, J.F., & Rivkin, S.G. (1998). Teachers, schools, and academic achievement. Retrieved October 27, 2005, from http://www.nber.org/papers/w6691.

Harman, A. E. (2001). National board for professional teaching standards' national teacher certification. Retrieved March 5, 2004, from www.ericfacility.net/extra/ericdigests/ ed460126.html (ERIC Document Reproduction Service No. ED460126)

Hawk, P., Coble, C.R., & Swanson, M. (1985). Certification: It does matter. *Journal of Teacher Education, 36*(3), 13-15.

Humphrey, D.C., Koppich, J.E., & Hough, H.J. (2004, November). *Sharing the wealth: National board certified teachers and the schools that need them most.* Menlo Park, CA: SRI International.

Johnston, R.C. (1999, May). Texas study links teacher certification, student success. *Education Week*, *18*, 19-20.

Jordan, H., Mendro, R., & Weerasinghe, D. (1997, July). Teacher effects on longitudinal student achievement. Paper presented at the Sixth Annual Evaluation Institute sponsored by CREATE, Indianapolis.

Kaplan, L.S. & Owings, W.A. (2002). The politics of teacher quality: implications for principals. *NASSP Bulletin*, *86*, 22-41.

King Rice, J. (2003).*Teacher quality: Understanding the effectiveness of teacher attributes.* Washington, DC: Economic Policy Institute.

Knoeppel, R.C. (2007) Resource adequacy, equity, and the right to learn: Access to a quality teacher in Kentucky. *Journal of Education Finance 32*(4), 422-442.

Krueger, A.B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics, 114*, 397-532.

Laczko-Kerr, I., & Berliner, D.C. (2002, September). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A

case of harmful public policy. *Education Policy Analysis Archives, 10*(37). Retrieved September 10, 2002, from http://epaa.asu.edu/epaa/v10n37/

Legislative Research Commission. (2006). Kentucky Revised Statute 157.360, *Base funding level—adjustment—enforcement of maximum class sizes—allotment of program funds.* Frankfort, KY: Kentucky General Assembly.

Loeb, S. (2001). Teacher quality. Its enhancement and potential for improving pupil achievement. In D.H. Monk, H.J. Walberg, & M.C. Wang (Eds.), *Improving educational productivity*. Information Age Publishing: Greenwich, CT.

Marzano, R.J., Pickering, D.J., & Pollock, J.E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: ASCD.

McColskey, W., Stronge, J.H., Ward, T.J., Tucker, P.D., Howard, B., Lewis, K., & Hindman. (2006). *A comparison of National Board Certified Teachers and non-National Board Certified Teachers: Is there a difference in teacher effectiveness and student achievement.*

Monk, D.H. (1989). The education production function; Its evolving role in policy analysis, *Educational Evaluation & Policy Analysis, 11,* 31-45.

Monk, D., & King. J. (1994). Multi-level teacher resource effects on pupil performance in secondary mathematics and science. In R.G. Eherenberg (Ed.), *Choices and consequences* (pp. 29-58). Ithaca, NY: ILR Press.

Murnane, R.J., & Phillips, B. (1981, March). What do effective teachers of inner-city schools have in common? *Social Science Research, 10*(1), 83-100.

National Board for Professional Teaching Standards (NBPTS). (n.d.a). History and facts. Arlington, VA: Author. Retrieved October 7, 2005, from www.nbpts.org/about/hist.cfm

National Board for Professional Teaching Standards (NBPTS). (n.d.b.). *About NBPTS*: *Five core propositions.* Retrieved October 7, 2005, from http://www.nbpts.org/about/coreprops.cfm

National Board for Professional Teaching Standards (NBPTS). (2004). *Q & A Questions and answers about National Board Certification.* Retrieved October 7, 2005, from NBPTS Web site: http://www.nbpts.org/

Nye, B., Hedges, L.V., & Konstantopoulos, S. (2000). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Education Evaluation and Policy Analysis 21*(2), 127-142.

O'Sullivan, R., Hudson, M., Orsini, M., Arter, J., Stiggins, R., & Iovacchini, L. (2005, June). *Student achievement and performance.* Chapel Hill, NC: University of North Carolina.

Odden, A., Borman, G., & Fermanich, M. (2004). Assessing teacher, classroom, and school effects, including fiscal effects. *Peabody Journal of Education, 79*(4), 4-32.

Pedhazur, E.J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. New York: Harcourt Brace College Publishers.

Pedhazur, E.J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). New York: Holt, Rinehart, and Winston.

Quality counts 2003. (2003). If I can't learn from you. . .ensuring a highly qualified teacher for every classroom, *Education Week, 22*(17), p. 7.

Rebell, M. A., & Wardenski, J. J. (2004). *Of course money matters: Why the arguments to the contrary never added up.* New York: Campaign for Fiscal Equity.

Reeves, D. B. (2006). *The learning leader: How to focus school improvement for better results.* Alexandria, VA: ASCD.

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

Rockoff, J.E. (2004). The impact of individual teachers on student achievement: Evidence from panel data, *American Economic Review, 94*(2), 247-252.

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 teachers. *Teachers College Record, 104*(8), 1525-1567.

Sanders, W.L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education, 14*(4), 329-339.

Sanders, W.L., Ashton, J.J., & Wright, S.P. (2005). *Comparison of the effects of NBPTS Certified Teachers with other teachers on the rate of student academic progress.*

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

Sanders, W.L., Wright, S.P., & Ross, S.M. (1999). *Value-added achievement results for two cohorts of Roots and Wings schools in Memphis: 1995-1998 outcomes.* Memphis, TN: Center for Research in Educational Policy, The University of Memphis.

Schwartz, A.E., & Steifel, L. (2004). Immigrants and the distribution of resources within an urban school district. *Educational Evaluation and Policy Analysis, 26*(4), 303-327.

Schwartz, A.E. & Zabel, J.E. (2005). The good, the bad, and the ugly: Measuring school effiiciency using school production functions. In S. Stiefel, A.E. Schwartz, R. Rubenstein, & J. Zabel (Eds.). *Measuring school performance and efficiency: Implications for practice and research*—American Education Finance Association 2005 yearbook. Larchmont, NY: Eye on Education. 37-66.

Sheskin, D. (2000). *Handbook of parametric and nonparametric statistical procedures* (2nd ed.). Boca Raton, FL: Chapman & Hall.

Smith, T.M., & Desimone, L.M. (2005). "Highly qualified" to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school mathematics, *Educational Evaluation and Policy Analysis, 27*(1), pp. 75-109.

Southeast Center for Teaching Quality (2004, April). Holding teachers accountable for value-added to student test performance. *Teaching Quality Research Matters.* Retrieved April 6, 2004, from http://www.teachingquality.org Web site.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* 3rd ed., Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Stevens, J. (1992). Applied multivariate statistics for the social sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates. {CHECK YEAR}

Stiefel, S., Schwartz, A.E., Rubenstein, R., & Zabel, J. (2005). Measuring school efficiency: What have we learned? In S. Stiefel, A.E. Schwartz, R. Rubenstein, & J. Zabel (Eds.). *Measuring school performance and efficiency: Implications for*

practice and research—American Education Finance Association 2005 yearbook. Larchmont, NY: Eye on Education.

Stone, J.E. (2004). *National Board for Professional Teaching Standards (NBPTS): Improving education at a snail's pace.* Retrieved February 28, 2005, from http://www.education-consumers.com/Urban_Inst1.asp .

Stronge, J.H., & Tucker, P.D. (2000). *Teacher evaluation and student achievement.* Washington, DC: NEA.

Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.) Boston: Pearson Education, Inc.

Tabachnick, B.G., & Fidell, L.S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins.

The Teaching Commission. (2004). *Teaching at risk: A call to action.* New York: The Teaching Commission, The Cuny Graduate Center.

Tobin, T.J., & Sugai, G.M. (1999). Using sixth-grade school records to predict school violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders, 7*(1), 37-54.

Topping, K.J., & Sanders, W.L. (2000). Teacher effectiveness and computer assessment of reading: Relating value added and learning information system data. *School Effectiveness and School Improvement, 11*(3), 305-337.

U.S. Department of Education. (n.d.). *No Child Left Behind (archived information): Who defines "highly qualified" teacher.* Retrieved July 2, 2007, from http://www.ed.gov/teachers/nclbguide/toolkit_pg23.html

Vandervoort, L.G., Amrein-Beardsley, A., & Berliner, D.C. (2004, September). National Board Certified Teachers and their students' achievement. *Education Policy Analysis Archives, 12*(46). Retrieved September 9, 2004, from http://www.epaa.asu.edu.epaa/v12n46/

Wiggins, G. (1993). *Assessing student performances: Exploring the purpose and limits of testing.* San Francisco: Jossey-Bass.

Wolfe, P. (2001). *Brain matters: Translating research into classroom practice.* Alexandria, VA: ASCD.

Word, E., Johnson, J., Bain, H.P., Fulton, B., Zaharies, J.B., Lintz, M.N., Achilles, C.M., Folger, J., & Breda, C. (1990). *Student teacher achievement ratio (STAR),*

Tennessee's K-3 class size study: Final summary report, 1985-1990. Nashville, TN: Tennessee State Department of Education.

Wright, S.P., Horn, S.P., & Sanders, W.L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57-67.

Wycoff, J. (2003, October). *Placing the preparation and recruitment of teachers in a labor market framework.* American Enterprise Institute for Public Policy Research, Progressive Policy Institute. Retrieved March 17, 2004, from http://www.ppionline.org

Figure 1
Theoretical Framework

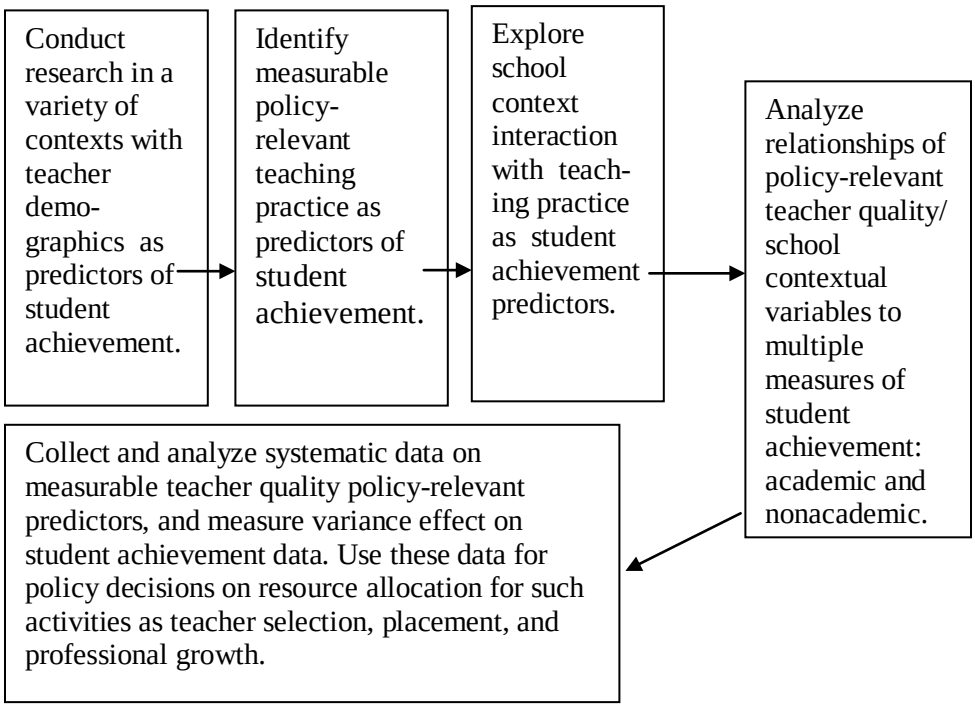

Figure 1. A conceptual framework for policy-relevant research on teacher quality: (a) analyze student achievement relationships to teacher demographics, (b) teaching practice, (c) school contextual variables that interact with teaching practice, and (d) verify these measurable policy-relevant teacher quality variables in a variety of school contexts with multiple measures of student achievement.

Table 1
Descriptive Statistics and Correlations for
2004 CATS Index and Measures of Teacher Quality

	N	Mean	Std. Dev	*r*	*p*
2004 CATS	339	75.73	10.40	1.000	-
% LEP	339	.91	2.03	-.010	.425
% Free/Reduced	339	46.47	19.60	-.481	.000**
% Special Ed	339	13.37	4.61	-.246	.000**
Assess Per Pupil	339	$375,102	$157.003	.128	.009**
% Major/Minor	339	97.45	7.16	.196	.000**
Avg. Years Exp.	339	11.91	2.40	.240	.000**
Student-Tchr Ratio	339	16.40	2.44	.064	.120
%NBCT	339	1.89	4.03	.230	.000**
% Masters	339	76.11	11.59	.116	.016*

** correlation is significant at the $p \leq .01$ level

* correlation is significant at the $p \leq .05$ level

Table 2
Multiple Regression

Model	Variables Entered	R	R Square	R Square Change	F Change	Sig F Change
1	% LEP % Free/Reduced % Special Ed	.483	.233	.233	33.998	.000
2	% LEP % Free/Reduced % Special Ed Assess Per Pupil	.484	.234	.000	.172	.678
3	% LEP % Free/Reduced % Special Ed Assess Per Pupil % Major/Minor	.514	.264	.030	13.638	.000
4	% LEP % Free/Reduced % Special Ed Assess Per Pupil % Major/Minor Avg. Years Exp.	.533	.284	.020	9.372	.002
5	% LEP % Free/Reduced % Special Ed Assess Per Pupil % Major/Minor Avg. Years Exp. Student-Tchr Ratio	.547	.299	.015	7.143	.009
6	% LEP % Free/Reduced % Special Ed Assess Per Pupil % Major/Minor Avg. Years Exp. Student-Tchr Ratio %NBCT	.560	.313	.014	6.707	.010

Table 2
Multiple Regression Continued

		Unstandardized Coefficients		Standardized Coefficients		
Model	Variables Entered	B	Std Error	Beta	Sig	Tolerance
1	Constant	88.312	1.654		.000	
	% LEP	.126	.246	.025	.608	.989
	% Free/Reduced	-.247	.028	-.465	.000	.799
	% Special Ed	-.093	.120	-.041	.441	.803
2	Constant	88.947	2.255		.000	
	% LEP	.162	.261	.032	.536	.885
	% Free/Reduced	-.251	.030	-.473	.000	.701
	% Special Ed	-.086	.122	-.038	.478	.790
	Assess Per Pupil	-1.5E-006	.000	-.022	.678	.809
3	Constant	63.670	7.193		.000	
	% LEP	.149	.256	.029	.562	.884
	% Free/Reduced	-.237	.030	-.447	.000	.690
	% Special Ed	-.141	.120	-.062	.242	.778
	Assess Per Pupil	-1.8E-007	.000	-.003	.959	.801
	% Major/Minor	.255	.069	.176	.000	.974
4	Constant	55.484	.7.591		.000	
	% LEP	.207	.253	.041	.414	.879
	% Free/Reduced	-.226	.030	-.427	.000	.681
	% Special Ed	-.103	.119	-.045	.391	.769
	Assess Per Pupil	5.7E-007	.000	-.009	.869	.800
	% Major/Minor	.253	.068	.174	.000	.974
	Avg. Years Exp.	.633	.207	.146	.002	.944
5	Constant	68.403	8.943		.000	
	% LEP	.125	.253	.024	.621	.866
	% Free/Reduced	-.246	.031	-.465	.000	.640
	% Special Ed	-.164	.121	-.073	.173	.741
	Assess Per Pupil	5.4E-007	.000	.008	.873	.800
	% Major/Minor	.235	.068	.162	.001	.964
	Avg. Years Exp.	.655	.205	.151	.002	.942
	Student-Tchr Ratio	-.585	.219	-.137	.008	.803
6	Constant	67.640	8.872		.000	
	% LEP	.071	.252	.014	.779	.860
	% Free/Reduced	-.230	.031	-.433	.000	.613
	% Special Ed	-.168	.119	-.074	.161	.741
	Assess Per Pupil	-2.9E-007	.000	-.004	.931	.799
	% Major/Minor	.225	.068	.155	.001	.961
	Avg. Years Exp.	.675	.204	.156	.001	.941
	Student-Tchr Ratio	-.578	.217	-.135	.008	.802
	%NBCT	.315	.122	.122	.001	.937