

ABSTRACT

Title of Dissertation: DOES CERTIFICATION OF ELEMENTARY SCHOOL TEACHERS MATTER? THE EFFECTS OF CERTIFICATION STATUS ON INSTRUCTIONAL PRACTICES AND ON THE MATHEMATICS AND READING ACHIEVEMENT OF FIRST GRADE PUBLIC SCHOOL STUDENTS

Richard J. Nyankori, Doctor of Philosophy, 2005

Dissertation directed by: Robert Croninger
Department of Education Policy and Leadership

This study examines teachers' certification status—emergency, standard, or advanced— as a predictor of teachers' instructional practices and of mathematics and reading of first grade public school students. The study is a secondary data analysis of the Early Childhood Longitudinal Survey (ECLS-K) and uses ordinary least squared regression as the primary statistical method. The chief finding is that certification, on its face, does not predict either the mathematics or reading achievement of first grade students when students' race, socioeconomic status, prior achievement, teachers' experience, and advanced degrees are controlled. The strongest predictors of first grade reading and mathematics achievement are students' prior achievement, SES, and race.

Certification status did have noteworthy indirect effects (i.e. OLS interaction terms) on both mathematics and reading achievement. In reading and mathematics, when emergency certification status was considered with end-of-kindergarten achievement, the results indicated that the students of teachers with emergency certification made fewer gains in reading achievement than the students of teachers with standard certification.

Similarly, in mathematics, when advanced certification status was considered with prior mathematics achievement, the results indicated that the achievement of students of teachers with advanced certification remained relatively unchanged.

Likewise, certification status did not directly predict the types of instructional practices that first grade teachers utilize in the classroom. Similarly, certification status had significant indirect effects on the examined instructional practice variables in mathematics. Emergency certified teachers who used number sense instruction decreased mathematics achievement scores.

The study concludes that the indirect effects of certification status on student achievement should signal educators that use emergency certified teachers may create inequities that result in diminished achievement for the most high need students. Therefore, the recommendations proposed encourage educators and policymakers to retool current certification practices and ensure that first grade students are taught by teachers with at least standard certification.

DOES CERTIFICATION OF ELEMENTARY SCHOOL TEACHERS MATTER?
THE EFFECTS OF CERTIFICATION STATUS ON INSTRUCTIONAL PRACTICES
AND ON THE MATHEMATICS AND READING ACHIEVEMENT OF FIRST
GRADE PUBLIC SCHOOL STUDENTS

by

Richard J. Nyankori

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2005

Advisory Committee:

Dr. Robert Croninger, Chair, Associate Professor
Dr. Jacqueline Cossentino, Assistant Professor
Dr. George Marx, Dean's Representative, Professor Emeritus
Dr. Hanne Mawhinney, Associate Professor
Dr. Jennifer King Rice, Associate Professor

©Copyright by
Richard J. Nyankori
2005

TABLE OF CONTENTS

List of Tables	v
Chapter I: Introduction.....	1
Defining Teacher Certification	2
Types of Teacher Certification	4
Early Childhood Education Certification	7
The Pathways to Certification.....	8
New Options for Teacher Certification	9
Two Central Purposes for Certifying Teachers: Regulation and Professionalization.....	11
Why Studying Teacher Certification Matters	13
Conceptual Framework: An Aggregated View of Certification Status	17
Research Questions.....	19
Conclusion and Overview of Chapters	21
Chapter II: Literature Review	23
Teacher Certification and Student Achievement	23
Level of Certification and Student Achievement	24
Subject-Specific Certification and Student Achievement	28
Comparisons of Alternately and Traditionally Certified Teachers.....	29
Teacher Certification and Instructional Practices.....	32
Empirical Approaches for Measuring Instructional Practices	34
Classroom Instructional Practices and Student Outcomes	36
Reform-Oriented Instructional Practices and Student Achievement.....	39
Reform-Oriented Instructional Practices Suggested for Teachers of Young Students.....	41
Chapter III: Methodology	45
Description of Data.....	45
Instruments.....	47
Direct Child Assessment.....	48
Spring First Grade Teacher Questionnaire	50
Spring Administrator Questionnaire	51
Variables	51

Dependent Variables.....	51
Independent Variable.....	52
Control Variables.....	53
Analytic Sample.....	58
Statistical Procedures.....	59
Chapter IV: Findings	60
Descriptive Analyses of Students, Teachers, Schools, and Classrooms.....	61
Student Descriptives	61
Teacher Descriptives.....	64
Classroom and School Descriptives	67
Summary	68
Principal Component Analyses of Identified Instructional Practices	69
Reading Instructional Practices	70
Mathematics Instructional Practices	73
Summary	76
Correlation Analysis of Teacher Instructional Practice Variables and Student Achievement Scores.....	77
Reading Instructional Practices	77
Mathematics Instructional Practices	77
Descriptive Analyses of Teachers' Use of the Identified Instructional Practices by Certification Status	79
Reading Instructional Practices	79
Mathematics Instructional Practices	80
Summary	81
Results of OLS Regression Analyses	82
Reading Scores	86
Mathematics Scores	88
Summary	90
Chapter V: Discussion and Conclusion	92
Review and Discussion of Major Findings.....	92
Examination of Research Questions	102
Research Question 1	102
Research Question 2	103

Research Question 3	103
Does Certification Matter?.....	104
Why Certification Doesn't Matter	105
Why Certification Matters	108
Summary	112
Revisiting the Literature	113
Teacher Certification and Student Achievement	114
Certification Status and Teachers' Instructional Practices	117
Limitations and Recommendations for Future Research.....	121
Conclusion	123
Appendix A: The National Center for Alternative Certification's Categorization of Alternate Route and Certification Programs	125
Appendix B: Summaries of Selected Research Grouped by Topic	127
References.....	143

LIST OF TABLES

Table 1.1. Definitions of Initial/Provisional/Probationary, Standard, and Advanced Certification and Certification Endorsements	5
Table 3.1. Description and Summary of Overall Achievement Variables ^a	52
Table 3.2. Description of Independent Variable ^a	53
Table 3.3. Description and Summary of Student Characteristic and Achievement Score Variables ^a	55
Table 3.4. Summary of Teacher Characteristic Control Variables ^a	56
Table 3.5. Summary of School and Classroom Characteristic Control Variables ^a	58
Table 4.1. Background Characteristics of First Grade Students in the Analytic Sample (N = 3,151) ^a	62
Table 4.2. Demographic and Education Characteristics of First Grade Teachers in the Analytic Sample (N = 3,151) ^a	65
Table 4.3. Characteristics of First Grade Teachers' Classrooms and Schools in the Analytic Sample (N = 3,151) ^a	68
Table 4.4. Preliminary Principal Component Analysis Results for Reading Instruction	71
Table 4.5. Final Principal Component Analysis Results for Reading Instruction	72
Table 4.6. Preliminary Principal Component Analysis Results for Mathematics Instruction	74
Table 4.7. Final Principal Component Analysis Results for Mathematics Instruction	75
Table 4.8. Summary of Correlation Coefficients for Reading Instructional Practice Variables and First Grade Reading Achievement (N = 3,151) ^a	78
Table 4.9. Summary of Correlation Coefficients for Mathematics Instructional Practice Variables and First Grade Mathematics Achievement (N = 3,151) ^a	78
Table 4.10. Descriptive Statistics for Reading Instructional Practice Variables by Teacher Certification Level (N = 3,151) ^a	80
Table 4.11. Descriptive Statistics for Mathematics Instructional Practice Variables by Teacher Certification Level (N = 3,151) ^a	81
Table 4.12. Summary of OLS Regression Analysis for Variables Predicting Average First Grade Reading IRT Scores (N = 3,151) ^a	84
Table 4.13. Summary of OLS Regression Analysis for Variables Predicting Average First Grade Mathematics IRT Scores (N = 3,151) ^a	85
Table B1. Research Comparing Teacher Certification Status and Student Achievement	127

Table B2. Studies Comparing Alternatively Certified Teachers With Traditionally Certified Teachers.....	130
Table B3. Studies Examining Certification Based on Teacher Grade Level or Subject Area.....	132
Table B4. Empirical Research Focused on the Relationships Between Classroom Practices and Student Achievement.....	135
Table B5. School Effectiveness Studies and Early Empirical Studies on Classroom Instruction and Student Outcomes.....	138
Table B6. Studies Focused on the Instructional Practices of Teachers of Young Children.....	140

CHAPTER I: INTRODUCTION

The commitment to ensure that all students reach high academic standards largely defines the current educational reform landscape. Most states are beyond the early stages of public school accountability and standards-based reforms and are moving closer to defining the resources or inputs needed to fulfill their promises. One resource, good teachers, has received considerable attention from policymakers and researchers alike. A growing body of empirical evidence strongly indicates that good teachers are critical to student achievement (Ferguson, 1998; Hanushek, Kain, & Rivkin, 1999; Goldhaber, Brewer, & Anderson, 1999; Wright, Horn, & Sanders, 1997).

Hanushek (1992) reported, for example, that students in a large urban district who had “good” teachers—those whose classes made large gains in achievement—gained nearly one and a half grade-level equivalents over the course of the year. Conversely, students assigned to classrooms with “bad” teachers realized gains of only half a year in a single academic year. Similarly, Sanders and Rivers (1996) found that “the effects of teachers on student achievement are both additive and cumulative with little evidence of compensatory effects” (p. 1). Illustrating Sanders and Rivers’s claim, a research team studying the mathematics achievement of elementary students in Dallas determined that students who were assigned to a highly effective teacher three years in a row had an average mathematics score in the 76th percentile. In contrast, students assigned to an ineffective teacher three years in a row had an average mathematics score in the 26th percentile (Jordan, Mendro, & Weerasinghe, 1997).

In support of these examples and others, Huang, Yi, and Haycock (2002) conclude that teachers “have the single greatest effect on student learning” (p. 1). Given that national education policy demands that schools show students are making “adequate yearly progress,” a teacher’s ability to affect student learning is critical. In response to this and other pressures, a chief challenge for policymakers, school officials, and researchers is to identify good teachers.

The sections that follow address how states use teacher certification to identify good teachers. Ultimately this study is concerned with understanding whether certification of elementary school teachers really matters and, thus, indicates quality—or might there be another way to identify good teachers for young students?

This discussion begins by attempting to define teacher certification and reveals the difficulty involved in articulating an exact definition of certification, given that its meaning varies considerably across contexts. The next section contains descriptions of the various types of certification—emergency, provisional, standard, and advanced professional—that states grant to public school teachers. Then, the pathways to certification and new certification options that a teacher can take to become certified are described. The focus here is on alternative certification, National Board Certification, and American Board Certification. The section concludes with a discussion of the two key purposes of certification: gatekeeping and professionalization.

Defining Teacher Certification

The mechanisms for identifying good teachers are limited. Goldhaber (2002) speculates that measurable teacher attributes—experience, degrees, teacher coursework, the teacher’s own examination scores, and certification—explain only 3% of the

differences in student learning. The other 97%, he argues, is attributable to “intangible aspects” of teacher quality, such as caring and enthusiasm that cannot be isolated or measured. The research evidence, however, shows that of measurable teacher attributes, verbal ability (Ehrenberg & Brewer, 1995; Greenwald, Hedges, & Laine, 1996; Hanushek, 1981, 1986), the prestige of the college or university the teacher attended (Ehrenberg & Brewer, 1994), and subject-matter competence correlate highly with student learning (Rowan, Chiang, & Miller, 1997; Shulman, 1987). The findings concerning the effects of other measurable teacher attributes on student learning are either inconsistent. Policymakers, nonetheless, continue to enact and back policy measures based on incomplete evidence and still expect to derive benefits.

One policy measure, the practice of certifying teachers, has been a prominent strategy for engendering teacher quality and is the key variable of interest in this study. All states have provisions for certifying or licensing teachers. The terms *teacher certification* and *teacher licensure* are used interchangeably throughout the literature to describe teacher credentialing, though some (Council of Chief State School Officers, 1992; Pyburn, 1990) have argued for differentiating between the two. Licensing, some argue, refers to a legal designation that protects the public from harm by an incompetent individual. Certification, on the other hand, refers to professional standards developed by members of the profession, which require more advanced expertise and experience than are typically required for licensure. Under this scheme, licensure may be a precursor to certification. The use of either term seems to depend chiefly on the statutory language a state uses to define the credentials needed to become a teacher. This study uses the term *certification* as inclusive of licensure.

Since states set the conditions and requirements for certification, what constitutes a teacher's certification varies greatly throughout the country, making it difficult to define the term accurately (Tryneski, 1997). Broadly defined, however, teacher certification refers to the process by which the state grants a permit to an individual to teach, as a means of guaranteeing the public that the person is qualified to practice in the profession (Lilly, 1992; Lortie, 1975; Wise, 2003). A teaching credential is usually granted after the completion of an accredited teacher preparation program, practice teaching, and the passing of a certification test.

Types of Teacher Certification

States typically offer several different types of teaching certificates with endorsements indicating specialized training. Table 1.1 provides definitions of the most common types of certification available to teachers in a state. Generally, certification denotes completion of coursework, satisfactory performance as judged by school officials, and time on the job (Kaye, 2002). Furthermore, teacher certification can be highly specialized, with designations for grade level, subject matter, specialized knowledge, or experience (McBrien & Brandt, 1997). For example, teachers of young children may be granted early childhood education certification, which signifies specialized training for the education of children 3 to 7 years old. The next section discusses early childhood education certification in greater depth.

Table 1.1. Definitions of Initial/Provisional/Probationary, Standard, and Advanced Certification and Certification Endorsements

Type of Certification	Definition
Initial/provisional/probationary certification	<p>Initial/provisional/probationary licenses are usually issued to a teacher who is beginning in the profession and has not met tenure requirements, if necessary. Teachers usually advance to the next level of certification after 2 or more years of satisfactory teaching experience. The teacher’s immediate supervisor (building administrator, district supervisor, department chair) usually makes the determination whether the provisional teacher’s performance has been satisfactory, usually based on observation of the teacher. Most initial certificates are issued for 3 or fewer years, with an option for renewal. (Sources: Kaye, 2002; Education Commission of the States [ECS], 2004)</p>
Standard certification	<p>Following their initial certification, teachers are usually granted standard certification. This level of certification denotes that the teacher is no longer a novice, usually has tenure, has 2 or more years of satisfactory teaching experience, and has completed additional professional development (either higher education coursework or in-service coursework). Many states have two or more types of standard certification. The difference between the two denotes that additional requirements (satisfactory time on the job, additional professional development) have been met for the higher level of certification. Most standard certificates are issues for 5 to 7 years, with an option for renewal that involves completion of additional professional development. (Sources: Kaye, 2002; ECS, 2004)</p>
Advanced certification	<p>Advanced certification, sometimes referred to as advanced professional certification, is the terminal certification status offered by most states. A key requirement for this certification is significant higher education coursework beyond a bachelor’s degree, an earned master’s, or a PhD. Although it is increasingly rare, some states grant advanced professional certification for a teacher’s lifetime. (Source: Kaye, 2002)</p>

Type of Certification	Definition
Endorsements or extensions	Usually issued in conjunction with either standard or advanced professional certification, endorsements or extensions can be added to a teacher's base certification. These endorsements usually specify that a teacher is "qualified" to (a) teach in a specific subject area (e.g., mathematics, social studies, or physical education); (b) teach a certain grade level (e.g., elementary, middle, or high school, or vocational); (c) teach certain types of students (e.g., special education students, English language learners, or young children); (d) or hold a nonclassroom instructionally related position such as that of librarian or technology specialist. Usually, endorsements are valid for the same period as the certificates to which they are attached. (Source: Kaye, 2002)

Though the practice is being phased out, primarily in response to No Child Left Behind (NCLB) mandates, many states issue emergency (or temporary) certificates that allow individuals to teach for a limited amount of time, usually a year, before having to complete the state's regular certification requirements. Under NCLB, states have been pressured to end the practice of issuing emergency certificates or certification waivers, because teachers possessing such certificates will not be considered to meet the "highly qualified" status by the 2005–2006 school year as required by NCLB (United States Department of Education, 2002). It is also important to note that in some states, provisional certificates serve the same purpose as emergency certification (Kaye, 2002). In other states, provisional or probationary certification denotes a teacher's "regular" initial certification, which he or she maintains prior to obtaining tenure. Some states assign alternatively certified teachers to an emergency status. In this study, emergency

certification refers to the practice of bypassing the regular certification criteria, including those established for alternatively certified teachers, to place a teacher in the classroom.

Emergency certificates are issued for a variety of reasons. The most common reasons include (a) shortages in critical subject-matter areas and high-need geographic regions (Rozycki, 1999), (b) teachers who are unable to meet one or more of the criteria required for initial or standard certification (Kaye, 2002), (c) inefficient recruitment practices (Levin & Quinn, 2004), and (d) a laissez faire attitude toward professionalism and strong certification requirements (Wise, 2003).

Early Childhood Education Certification

Since young students are a primary focus of this study, the certification granted to teachers of these students is of interest. Early childhood education (ECE) certification is relatively new, and like other types of certification its meaning and use varies from state to state. Generally, ECE certification applies to teachers of students up to age 8, which usually includes public school kindergarteners and sometimes first graders (National Institute for Early Education Research [NIEER], 2004). The ECE knowledge base emphasizes developmental psychology, developmentally appropriate pedagogy, and play and socialization, and engages the child's family in learning (Charlesworth, 1998; National Association for the Education of Young Children [NAEYC], 2004).

Initially ECE certification was offered through community colleges to preschool teachers and caregivers as a prerequisite to the licensing of daycare programs, and largely focused on granting child development associate (CDA) status (Bowman, Donovan, & Burns, 2001). Many states use the CDA designation as a basic requirement for preschool teachers. Influential groups such as the NAEYC have specified standards for teachers of

young children that have subsequently been adopted by state certification agencies. Additionally, several states are pushing for universal kindergarten and prekindergarten programs, and ECE certification is seen as one strategy for improving teacher quality and ensuring that expenditure is productive (Ackerman, 2004). Further, policymakers raised the education and credential requirements for childcare workers (Head Start Act, 1998). Following the increasing interest in professional training for childcare workers working with young children, four-year colleges and universities began to offer programs of study in ECE that were based on both the care and instruction of young children and developmental psychology (Bowman, Donovan, & Burns, 2001). Currently, 48 states have an ECE certification option and seven states require ECE certification for first grade teachers (NIEER, 2004). The effects of ECE certification on student achievement are discussed in Chapter 2.

The Pathways to Certification

Teacher certificates are also differentiated by the path taken to certification. Darling-Hammond (1990) distinguishes between “alternate routes” and “alternative certification.” Alternate routes maintain the same certification requirements as those met by traditionally prepared teachers but offer different options for earning the certificate. Alternative certification usually entails the state altering the established rules for certification. Generally, alternatively certified candidates must have a four-year degree (though it is usually not in education) from an accredited college or university, complete a criminal background check, and pass a certification test for initial certification while completing additional requirements (e.g., pedagogical coursework, satisfactory

observation, or mentoring) concurrently with full-time teaching responsibilities (Feistritzer & Chester, 2000).

As is the case with traditional certification, many varieties of alternative certification exist. Feistritzer (2004) identifies 11 classes of alternative certification routes (see Appendix A for descriptions). Generally, these certification classes differ based on candidates' immediate qualifications, the amount and type of coursework or professional development required to obtain standard certification, and emergency certification needs. The most frequently used type of alternative route is Class D, the college- or university-based program (Feistritzer, 2004).

The efficacy of alternative certification is hotly debated among researchers, practitioners, and policymakers. Detractors decry it as a harmful practice damaging both the profession and the students taught by such teachers (Darling-Hammond, Berry, & Thoreson, 2001). Proponents claim the contrary, arguing that the practice allows for innovation in the preparation and recruitment of talented individuals who otherwise might not have considered teaching (Hess, 2001; Walsh, 2001).

New Options for Teacher Certification

In addition to certification usually earned in conjunction with college or university study, other entities are expanding certification options—endorsed by the state—for teachers. To date, the National Board of Professional Teaching Standards (NBPTS) and the American Board for Certification of Teacher Excellence (ABCTE) offer two significantly different certification options. The NBPTS offers National Board Certification to experienced teachers through a portfolio assessment that documents aspects of a teacher's performance and knowledge. Many states recognize National

Board Certification and reward recipients with increased compensation, recognition, choice teaching assignments, and career advancement (Stone, 2002). Bond, Jaeger, Smith, and Hattie (2000) examined the construct and consequential validity of National Board Certification and determined that NBPTS-certified teachers differed significantly from non-certified teachers in terms of student achievement. Likewise, Goldhaber and Anthony (2004) report that NBPTS is successful in identifying effective teachers among its applicants. They note, however, that these teachers were more effective than their counterparts even prior to seeking National Board Certification. Nevertheless, the researchers qualify these results by stating that the magnitude and statistical significance of the observed “NBPTS-effect” varies greatly by subject matter and grade level. Others (Podgursky, 2001; Stone, 2002) assert that National Board Certification is no more useful than conventional certification methods in identifying teachers who positively affect student achievement.

ABCTE works with states to grant initial certification to candidates, primarily on the basis of passing a computer-based assessment of their subject area and professional teaching knowledge. In addition to earning a passing score, which is set by the state, candidates must possess a bachelor’s degree and complete a criminal background check. ABCTE markets its certification specifically to career changers and liberal arts college graduates who would not consider teaching a viable option if they were required to attain certification through additional higher education coursework (Holland, 2004). So far, three states—Florida, Idaho, and Pennsylvania—accept this certification option. Recipients in each of these states are considered “highly qualified” under NCLB. To date, 11 teachers have received certification from ABCTE (American Board for Certification of Teacher Excellence, 2004). Due to the small sample and newness of

ABCTE certification, no research concerning the efficacy of this approach has been generated yet. ABCTE, however, received \$35 million from the federal government to further develop the content standards on which its assessments are based, to create assessments in additional subject areas, and to conduct a longitudinal study of the effects of ABCTE-certified teachers on student achievement. Independent reviews of ABCTE certification should be conducted as well. Even so, the outcry over ABCTE's "test and certify" or "click and certify" strategy has drawn harsh criticism from traditional teacher advocates such as the National Education Association, the American Federation of Teachers, the National Council for Accreditation of Teacher Education (NCATE), and the American Association of Colleges for Teacher Education (AACTE), whose collective control and influence over the way teachers are prepared and certified is potent (Imig, 2003).

*Two Central Purposes for Certifying Teachers:
Regulation and Professionalization*

Beyond indicating that a candidate possesses the necessary minimum credentials to qualify for a teaching position, the certification of public school teachers serves two central purposes—gatekeeping and professionalization (Dingwall & Fenn, 1987; Rice, 2003). Certification is both a primary gatekeeper regulating the flow of candidates into teaching and a tool used for professionalizing the teacher workforce (Rice, 2003). The gatekeeping function allows states to regulate who enters the workforce and to stipulate what qualifications candidates need in order to enter. In cases where demand exceeds supply, for example, states may ease certain certification requirements (or simply misassign teachers by assigning them to courses outside of their expertise) to meet its

needs even though doing so secures very few teachers (Ng, 2003). Additionally, states that experience staffing difficulties in high-need subject areas such as upper-level mathematics, special education, and English as a second language have changed certification requirements, incentives, and recruitment strategies to increase the pool of candidates in those areas (Recruiting New Teachers, 2000). Similarly, states may ease certification requirements in order to diversify the teaching pool in terms of gender, age, or race (Houston, Marshall, & McDavid, 1993; Shen, 1998).

Second, as a tool for professionalizing the teacher workforce, certification is a mechanism for articulating and exacting higher standards for teachers (Bacharach, 1990; Brown, 1995; Clifford, 1989; Darling-Hammond, Wise, & Klein, 1995/1999; Holmes Group, 1986). Efforts such as those by the Interstate New Teacher Assessment and Support Consortium (INTASC), NBPTS, NCATE, and a host of other professional teacher organizations have worked to define what teachers should know and be able to do. Though setting standards begins to standardize teachers' professional knowledge, Goldhaber and Anthony (2004) note that "there is considerable controversy about how teachers can actually achieve and demonstrate mastery of [these standards]" (p. 6). Simply setting standards and certifying teachers on the basis of meeting them may be a necessary but insufficient step toward professionalization goals. Alternatively, attending to other related issues such as improving working conditions and job satisfaction (Brandt, 1993; Ingersoll, 2001b), establishing professional learning communities (Kruse, Seashore Louis, & Bryk, 1994), and addressing compensation (Odden & Kelley, 2002) may be as likely as certification to enhance professionalization efforts.

In sum, the value of certifying teachers is far from straightforward. The proposition that certified teachers are good teachers demands further investigation,

clarification, and elaboration. The difficulty in defining certification, the variability of certification requirements across contexts, the different available routes to certification, and the multiple purposes of certification make studying its effects problematic. Researchers, policymakers, practitioners, and the public, nonetheless, express great interest in the subject. The next section provides a specific rationale for why studying certification matters, given these challenges.

Why Studying Teacher Certification Matters

Despite the uncertainty about the appropriateness of certification, the outpouring of fiscal, human, and political resources to ensure that every public school classroom has certified teachers has not waned. A number of influential interest groups have called for the strengthening of teacher certification requirements. The National Commission on Teaching for America's Future (NCTAF; 1996) recommended that policymakers and educators strengthen licensing requirements and end the practice of alternative certification (Darling-Hammond, Berry, & Thoreson, 2001; NCTAF, 1996). The president of NCATE, which offers accreditation to college- and university-based teacher education programs, called for making certification (tied to the group's criteria) a key driver in overhauling teacher education programs (Wise, 2003).

Moreover, NCLB requires states receiving Title I funds to have a "highly qualified" teacher in every public school classroom by the 2005–2006 school year. NCLB defines a "highly qualified" teacher as one who is fully licensed or certified by the state, demonstrates subject-matter competency, and has not had any licensure or certification requirements waived on an emergency, temporary, or provisional basis

(United States Department of Education, 2002).¹ Arguably, the “highly qualified” status defined by the United States Department of Education could serve as a low rather than a high bar for identifying teacher quality. Moreover, observers of states’ compliance with the “highly qualified” teacher mandate report that few states will substantively meet this goal, and only a handful of states have acted within the spirit of the law to meet it.

Rather, they argue, many states have misreported their progress toward the goal, several states have distorted their status, and some have refused to provide the data at all (Huang, Yi, & Haycock, 2002).

Despite the appealing rhetoric of placing a highly qualified teacher in every public school classroom, many (Ballou & Podgursky, 2000; Hess, 2001; Walsh, 2001) view the typical certification requirements as unnecessary barriers that may dissuade the best candidates from pursuing teaching jobs. Complicating matters, the “highly qualified” directive comes at a time when finding and keeping good teachers is an overwhelming task for many public schools and districts. The “graying” of the nation’s teacher workforce (Ingersoll & Smith, 2003), pronounced teacher shortages in certain subject areas (Hardy, 1998; Wayne, 2000), class size reduction efforts (Jepsen & Rivikin, 2002), and teacher dissatisfaction with the workplace (Billingsley, 1993; Brownell & Smith, 1992; Morvant, Gersten, Gillman, Keating, & Blake, 1995) represent only a few of the challenges involved in recruiting, selecting, retaining, and training highly qualified teachers to satisfy the mandate.

¹ The teacher quality provisions of No Child Left Behind (NCLB), in which the “highly qualified” teacher mandate originates, are the principal area of interest for this study. Other aspects of NCLB’s teacher quality provisions include ensuring that teachers have subject-matter mastery, creating high standards for paraprofessionals, developing mechanisms for tracking and disclosing information on teacher qualifications, and promoting ongoing professional development for teachers. There are also provisions for investing in teacher recruitment and reforming teacher certification processes. Additionally, NCLB holds

Although numerous studies have attempted to sort out the relationship between teacher characteristics, such as certification status, and desired educational outcomes, the present investigation is important for three reasons. First, the policy relevance of certification makes the study of interest. The terms of certification are largely controlled and driven by policymakers. Though some states use professional teaching standards boards² to provide input on certification policy, their work is more symbolic than substantive (Kaye, 2002). Certification regulators can shape policy for a variety of reasons, and the resulting requirements may or may not be beneficial to student learning, or even feasible.

Second, the unequal distribution of qualified teachers has been well documented (Haycock, 2002/2003; Ingersoll & Gruber, 1996). Students who most need the best teachers are least likely to be taught by them. These students tend to be ethnic or language minorities, or both, have low socioeconomic status, and attend schools in rural areas and large urban districts (NCTAF, 1996). Haycock (2002/2003) reports that on several dimensions of teacher qualifications—certification status, experience, subject-matter competence, teachers' exam performance, and classroom effectiveness—students in high-poverty schools are more likely to be taught by less qualified teachers. For example, she reports that 30% of core academic courses are taught by uncertified teachers in high-poverty secondary schools, compared to 17% in low-poverty schools (Haycock, 2002/2003). Teachers who performed poorly on licensing exams, a key part of

states accountable for ensuring that students make “adequate yearly progress” toward state-defined academic standards.

² Fourteen states have professional teacher standards boards or commissions that contribute to teacher credentialing policies but do not have the authority to issue teacher certificates outright. Rather, these boards, whose members are appointed by a state's governor or state superintendent of schools, have more of an advisory role for policy elites.

teacher certification, and college entrance tests are more likely to teach in high-poverty schools. Twenty percent of inexperienced teachers (i.e. teachers in their first few years of teaching) learn to teach in high-poverty schools, compared to 11% in low-poverty schools. Likewise, Ingersoll and Gruber (1996) determined that students in high-poverty schools were more likely to be taught by an out-of-field teacher than were students in low-poverty schools. Although this evidence is compelling, it is not clear whether equalizing the qualifications of teachers across student populations would effectively address inequities in student outcomes.

Third, this study focuses on the early elementary grades and the instructional classroom practices of early-grade teachers, which have not been adequately studied with respect to certification. Much of the empirical research on teacher certification has used data focused specifically on high school teachers. Where elementary grades are considered, the data are typically from smaller-scale research studies and program evaluations with limited validity. Moreover, little empirical research focuses on whether teachers' use of various instructional activities differs by certification status.

In summary, the use of certification to identify good teachers is a salient research topic. A critical aspect of this topic, however, rests on understanding the extent to which certification is a good indicator of teachers' ability to effect gains in student learning. The discussion above highlighted several of the complexities involved in studying this issue. First, although certification is used as a qualification for teaching in every state, how certification is defined and the path to earning it varies greatly across states and within them. Second, many types of certification exist. Teachers can earn different types of

certification based on subject area, grade level, experience, and additional study.

Teachers can also earn advanced certification either from the state or through National Board Certification. Third, certification serves two key purposes: professionalization and gatekeeping. Some view certification as a strategy for increasing the professional status of teachers, while others view it as way to control the flow of individuals who can enter the profession. Fourth, the issue of whether certification affects student learning continues to be important in understanding the efficacy of using it to identify good teachers.

This study utilizes large-scale survey research and secondary data analysis to examine whether public school first grade teachers of differing certification statuses—emergency, standard, and advanced—impact student achievement in reading and mathematics differently and vary in their use of classroom practices. Such an inquiry will help determine whether certification matters for teachers of first grade students. Before delving into the specific research questions addressed in this study, consider first the conceptual framework used to ground this analysis.

Conceptual Framework: An Aggregated View of Certification Status

Conceptualizing the effects of certification status on outcome variables in this study necessitates viewing certification as an aggregation of its component parts rather than the converse. A *disaggregated* approach to certification lends itself to examining the effects of each certification component—degree, experience, type of preparation program, coursework—on an outcome variable. Given that considerable variability exists in terms of the specific components that comprise certification in a given state, this approach is well suited for studies in which the certification components are known. This

approach also makes sense in helping to define the components that add value to the notion of certification generally. Future studies using the data set may wish to restrict the sample to isolate specific states that share similar certification requirements or focus on only one state.

Alternatively, the *aggregation* approach views certification as an amalgamation of the various components that comprise it. Instead of looking at the effects of each component, certification is accepted on a prima facie basis. Since policymakers and the public largely treat certification as a one-dimensional phenomenon rather than a multifaceted one, it is reasonable to view the value of certification on its face. That is, the study uses teachers' self-reports of their certification status (emergency, standard, and advanced). This approach is useful for studies, such as this one, that consider certification status using a nationally representative sample in which the components of certification are known to vary.

Certification status is one means of understanding teacher quality, but one with a questionable theory of action. This theory of action presupposes that as a teacher's level of certification increases so does the quality of their teaching. Therefore, one might expect that teachers of differing certifications statuses do something different in the classroom—that is, one would expect teachers with the highest levels of certification to utilize instructional practices that increase student achievement; and conversely emergency teachers' practices would be detrimental to student achievement. The focus on teacher performance is a complementary approach for understanding teacher quality in terms of not only credentials but also what they do and the effectiveness of what they do.

The aggregation approach does have limitations. First, although the approach acknowledges the variability question, it does not resolve the issues associated with the

different certification regulations found between states. Chief among these issues is the degree to which the findings based on teachers' self-reports of certification status are generalizable to the population from which the sample is drawn. Since there is an ordinal nature to the types of certification status used in this study, we know that teachers' responses show differentiation between types. For example, teachers indicating that they have standard certification means that they view their certification as higher than provisional but lower than advanced. Second, the findings generated can inform only broad notions about the effects of certification status on the achievement of young students. Accordingly, the disaggregated approach is best suited to developing a more nuanced understanding of "what" about holding certification matters for teachers of young students. For example, what types of subject-matter preparation, certification examination, and professional experience should comprise elementary certification?

Research Questions

This study addresses three specific research questions, which are outlined below. In the study, certification status is the primary independent variable, while classroom instructional activities and student achievement in mathematics and reading are the chief dependent variables. The first question seeks to describe how first grade public school teachers, their schools, and their students vary by certification status. The second question examines whether a teacher's certification status predicts the use of various classroom instructional activities. The third question investigates whether certification status predicts students' mathematics and reading achievement. These questions are explained more fully below.

Question 1: How do the characteristics of first grade public school teachers, their schools, and their students vary by certification status?

This question examines the school-, teacher-, and student-level characteristics of first grade public school teachers by certification status. The goal of this investigation is to provide a broad description of how teachers, their students, and their schools vary by certification status. Specifically, the study explores the following school-level characteristics: class size, percentage of ethnic minorities, and receipt of Title I funds. The study investigates the following teacher characteristics: demographic characteristics (age and race), years of teaching experience, and teacher degrees. Finally, the following student characteristics are examined: demographic characteristics (age, race, and gender), socioeconomic status (SES), and academic achievement.

Question 2: Does a teacher's certification status predict the frequency of use of various classroom instructional activities in reading and mathematics?

One primary assumption of certification is that it represents a teacher's professional knowledge base. Teachers exercise this professional knowledge, in part, by using instructional techniques that such knowledge makes available to them. This question seeks to determine whether a teacher's certification status predicts the types of classroom instructional activities used for mathematics and reading instruction and whether their chosen practices are related to increasing student achievement. Specifically, the study examines the following reading and language arts instructional practice variables: (a) letter sense instruction, (b) fluency and comprehension instruction, and (c) student-centered literacy instruction. In mathematics, the study explores the following instructional practice variables: (a) number sense instruction, (b) computation instruction, and (c) student-centered mathematics instruction.

Question 3: Does a teacher's certification status predict how students will perform in mathematics or reading, or both?

Student achievement figures prominently in the current education reform agenda. A student's early literacy and mathematics proficiency are important to his or her future academic success. This question seeks to understand whether a teacher's certification status predicts the performance of students in his or her classroom.

Conclusion and Overview of Chapters

The goal of Chapter 1 was threefold. First, because certification is complex, the definitions, types, options, and routes to certification vary greatly; the discussion above aimed to address these complexities. Second, the chapter aimed to show that teachers' certification status as a predictor of their classroom practices and their students' achievement is an important issue to study. The policy manipulability of certification, the need to equally distribute good teachers across student populations, and the contribution to the literature of certification status among the teachers of young children were cited as reasons why studying certification is important. Finally, the chapter articulated a framework for viewing certification in the context of a nationally representative sample of first grade teachers and specified three key research questions to guide the study.

Chapter 2 reviews the related literature. Specifically, three major bodies of literature are described: (a) teacher certification and student achievement, (b) teacher certification and classroom instructional practices, and (c) instructional practices of early-grade teachers. In the first area, the focus is on understanding what others have found with regard to the relationship between the type of certification a teacher holds and

subsequent student achievement, as well as grade levels and subject areas in which certification seems to matter. The value of early childhood education certification is also examined. The second section relates to patterns in classroom instructional practices that differ by certification status. The third and final area of research focuses on the instructional practices of early-grade teachers.

Chapter 3 presents the methodology used to address the study's research questions. The chapter begins by describing the Early Childhood Longitudinal Survey–Kindergarten Cohort (ECLS-K) database and instruments, followed by descriptions of the key dependent variables (i.e. mathematics and reading scores and classroom activities/skill areas) and independent variable (i.e. certification status). The chapter concludes with a description of the statistical models and procedures used in the analysis.

Chapter 4 presents the findings from the analysis. First, findings from the descriptive analyses of student-, teacher-, and school-level characteristics are given with respect to certification status—emergency, standard, and advanced. Second, the results of the principal component analysis (PCA) are provided. The findings from a correlation analysis of the instructional practice variables and mathematics and reading item response theory (IRT) scores follow the PCA. The analysis concludes with findings from two ordinary least squares (OLS) regression models (one in which the reading IRT score is the dependent variable and the other in which the mathematics IRT score is the dependent variable) to identify the predictors of mathematics and reading achievement.

Chapter 5 reviews and discusses eight key findings derived from the analysis. Then the findings are contextualized in the literature reviewed in Chapter 2. The chapter concludes with a presentation of the study's limitations and suggested future areas for research.

CHAPTER II: LITERATURE REVIEW

This chapter focuses on three major bodies of literature: (a) teacher certification and student achievement, (b) teacher certification and classroom instructional practices, and (c) empirical approaches for measuring the classroom instructional practices. The first section focuses on understanding what others have found with regard to the relationship between the type of certification a teacher holds and subsequent student achievement, as well as what others have found regarding grade levels and subject areas in which certification seems to matter most. The value of early childhood education certification is also examined. The second section reviews literature related to patterns in classroom instructional practices that differ by certification status. The final section focuses more generally on the instructional practices of early-grade teachers, including what others have found about the relationship between practice and student achievement.

Teacher Certification and Student Achievement

Dozens of studies have been conducted in an attempt to understand how a teacher's certification status is related to student achievement. The literature reviewed here is divided into three subcategories: (a) studies that focus on level of certification—advanced, standard, and emergency—and student achievement, (b) studies that focus on subject-specific certification and student achievement, and (c) studies that compare the achievement of students taught by teachers with traditional certifications and teachers with alternative certifications. These categories emerged from a simple grouping strategy undertaken in preparation for the current study in which articles were coded by their

major findings (see Appendix B, Tables 1.1B–1.6B, for the tables summarizing the major findings, methodology, and sample sizes of each of the studies). Overall, the results of this review reveal that a teacher’s level of certification shows little clear impact on student performance, and that highly specialized subject-specific certifications revealed only a positive effect for high school science and mathematics achievement. The reviewed research also contains mixed findings regarding the effects of alternatively certified teachers as compared to traditionally certified teachers.

Level of Certification and Student Achievement

Level of certification in this study is constructed as advanced, standard, and emergency. In the literature, many of the studies construct certification as a dummy variable in which teachers are recorded as having certification (i.e. standard or advanced certification) or not having certification (i.e. private school, temporary/provisional, or alternative certification). The construction of these certification categories is problematic, as indicated earlier. Given this tendency in the literature, the studies are divided into two groups. The first group of studies focuses on what others have found about the effects of standard and advanced certification on student achievement, while the other group of studies focuses on what others have found about the effects of emergency certified teachers. Although there is potential overlap in these groupings, each represents a dominant research design in the literature (the potential benefits of full certification and the potential hazards of emergency certification).

Advanced and Standard Certification. Five studies are reviewed below. Two studies utilized data from the National Assessment of Educational Progress (NAEP) but reported conflicting findings regarding the effects of full certification status (advanced or

standard compared to others). A third study failed to find a relationship between level of certification and student achievement. The fourth study provides evidence that students taught by fully certified teachers—standard or higher—scored higher on mathematics exams. The final study, a meta-analysis, reports that the empirical literature fails to find a clear and consistent connection between level of certification and student achievement—a conclusion reached by other reviews of the literature (see Rice, 2003).

Specifically, the *1990 Science Report Card* by the National Center for Education Statistics (NCES) analyzed national data from the National Assessment of Educational Progress (NAEP) and determined that none of the following teacher characteristics were statistically associated with student achievement: experience, certification level (either advanced or standard), master's degree, and coursework in the subject matter. Likewise, Goldhaber and Brewer's (1998) analysis of NELS:88 data, which controlled for race, SES, prior achievement and linked student records to teachers, suggested that teacher certification (at any level), teacher experience, and possessing a master's degree were not related to higher test scores for 10th grade students.

In contrast, Darling-Hammond's (1999) state-level analysis of NAEP utilized multiple regression, in which poverty and language status were controlled, and argued that certification and teacher preparation are the strongest correlates of student achievement in mathematics and reading. Though she acknowledged the threat and likelihood of aggregation bias (the actual analysis examined average state achievement scores and percentages of certified teachers), she dismissed the concerns by reasoning that the data are useful for state-level policy. Felter (1999) examined the effects of certification status at the student level. In a correlation and multivariate regression analysis of 797 California high school students, controlling for SES and race, Felter

found that students of teachers with full certification (i.e. standard or advanced) scored higher on mathematics exams than students with less than full certification.

Walsh (2001), however, called into question the findings of a number of studies, including that by Darling-Hammond (1999). In her review of nearly 150 published studies and papers related to certification, Walsh concludes that the empirical record is inconclusive, at best, with respect to the relationship between certification status and teacher effectiveness.

Emergency Certification. With respect to emergency certification or “undercertified” teachers, the evidence from the four studies reviewed present conflicting results as well. A meta-analysis by Qu and Becker (2003) shows that among various levels and types of certification, emergency teachers are generally found to be of the lowest quality. An empirical study by Laczko-Kerr and Berliner (2002) supported the conclusions drawn by Qu and Becker. However, a study by Goldhaber and Brewer (2000) found that emergency certified teachers performed similarly to standard certified teachers in high school mathematics. The findings of this study were contested by Darling-Hammond, Berry, and Thoreson (2001) who questioned the methodology in the Goldhaber and Brewer study. These studies are described in greater depth below.

Qu and Becker’s (2003) conducted a meta-analysis of 24 studies, using clear and reasonable selection criteria to identify high-quality studies that examine the effects of emergency certified teachers and traditionally certified teachers. Emergency certified teachers did less well than traditionally certified teachers on nearly all the outcomes (i.e. student achievement, teacher performance, and personality-like measures) encompassed by the studies. They also did somewhat less well compared to alternatively certified

teachers, though the latter “performed equivalently” to traditionally certified teachers on many outcomes.

Laczko-Kerr and Berliner (2002) made a similar determination in their study of “undercertified” teachers (i.e. those with emergency, alternative, provisional, and temporary certification) and fully certified teachers (i.e. those with standard and advanced professional certification). The researchers used a match-pairing design of 293 teachers in elementary schools in Arizona, and SAT-9 (Stanford Achievement Test, Ninth Edition) scores in mathematics, language arts, and reading as student achievement measures. In all areas, fully undercertified teachers performed less well than traditionally certified teachers. The study did not attempt to determine if there were important differences in performance of teachers with different types of “undercertification” credentials.

Goldhaber and Brewer (2000) provide a contrary argument in their econometric models that utilized the NELS:88 data. The researchers determined that students whose teachers held standard certification did better in mathematics than their peers whose teachers held private school certification or were not certified in the subject area. However, they found that students whose teachers held emergency certification did no worse than students whose teachers had standard certification. Darling-Hammond, Berry, and Thoreson (2001) disputed Goldhaber and Brewer’s findings that emergency certified teachers are as good as standard certified teachers. They replicated Goldhaber and Brewer’s (2000) study using the same data set (NELS:88) but concluded that the emergency label is a misnomer in NELS because the teachers with emergency certification in Goldhaber and Brewer’s sample actually resemble teachers with standard certification, thus discrediting their findings.

Subject-Specific Certification and Student Achievement

The literature reviewed in this section describes the relationships others have found with respect to subject-area and grade-level certification. Rice's (2003) review of the literature on several areas of teacher quality concluded, in part, that teacher certification seemed to matter most in the areas of high school mathematics and science and at the high school level in general. Certification seemed less a predictor of student achievement in other subject areas or at any other grade level.

Hawk, Coble, and Swanson (1985) conducted a match-paired analysis of 36 middle and high school teachers and found that students (n=826) of secondary teachers who were certified in mathematics had higher mathematics test scores than their counterparts taught by teachers not certified in mathematics. Goldhaber and Brewer's (1996) multivariate analysis of the NELS:88 data found that students of secondary school mathematics teachers who were certified in mathematics or who had majored in mathematics or earned a master's degree in mathematics had higher test scores than their peers whose teachers did not meet these criteria. In subsequent studies (Goldhaber & Brewer, 1997, 2000), the researchers confirmed this initial finding and determined that certification has a small but significant effect on student achievement in high school science.

Though certification seemed to matter in mathematics and science, the size of the effect appeared to be quite small. Mandeville and Liu's (1997) match-paired study of 9,000 seventh grade students in 33 schools found that students whose teachers had the highest level of preparation (i.e. advanced certification in mathematics) had only slightly improved scores on three measures of mathematics achievement (i.e. mathematical

reasoning, critical thinking, and computation) when compared to their peers whose teachers held elementary certification. The study, which used, multiple regression analysis controlled for SES, urbanicity, race, and school size but did not control for prior achievement in There is mixed evidence, however, that subject-specific certification matters at the elementary school level. With respect to elementary teachers, Rowan, Correnti, and Miller (2002) used a hierarchical linear growth model to analyze data from *Prospects: The Congressionally Mandated Study of Educational Growth and Opportunity*. They found that for elementary teachers, subject-specific certification was not related to increased student achievement in mathematics and reading. Nonetheless, another empirical study suggests that 1st grade students taught by teachers with an elementary education certification do better in the classroom than students taught by teachers with an early childhood education certification (Croninger, Rice, Rathbun & Nishio, in press). In short, the empirical evidence indicates that high school mathematics teachers with mathematics or science certification and middle school teachers with secondary certification had a positive effect on student achievement in mathematics. The evidence does not support any statistically significant effect of subject-matter certification for elementary school teachers in reading and mathematics. However, there is some evidence that elementary school certifications matters in the early grades.

Comparisons of Alternatively and Traditionally Certified Teachers

Though the present study does not focus on alternatively certified teachers specifically (the sample contains too few of them), the literature related to alternatively certified teachers highlights potentially important differences between teachers who hold traditional certification and those who do not. Generally the research revealed, as

discussed by Qu and Becker (2003), noted above, that the performance of alternatively and traditionally certified teachers was nearly equivalent.

First, supervisor reports of alternatively certified teachers were positive and generally higher than traditionally certified teachers. Lutz and Hutton's (1989) study of 100 alternatively certified teachers found that supervisors and mentor teachers rated alternatively certified teachers higher than first-year teachers with standard certification. Similarly, in Ashton's (1996) review of the research, he reported evidence that alternatively certified teachers received higher supervisor ratings than uncertified teachers.

In terms of principal ratings of alternatively certified teachers, Lutz and Hutton (1989) surveyed principals of nearly 100 alternatively certified teachers in Texas and found that principals and mentor teachers, across all grade levels, rated alternatively certified teachers as high or higher than first-year teachers with standard (i.e., initial) certification. Ashton (1996) found a similar pattern in his review of the literature described above. With respect to TFA teachers in particular, Kane, Parsons, and Associates (2005), in a report commissioned by TFA, found that 70% of the principals surveyed reported that TFA teachers were as good as or more effective than other first-year teachers, and 95% of principals said that they would hire another TFA teacher. Second, the performance of alternatively certified teachers when compared to teachers with other types or levels of certification appeared comparable and in some instances higher. For example, The students of teachers certified through the nation's largest and best-known alternatively certified recruitment program, Teach for America (TFA), performed better in mathematics and nearly the same in reading when compared to students of teachers with traditional certification (Decker, Mayer, & Glazerman, 2004).

This study is discussed in depth below. Brown, Edington, Spencer, and Tinafero (1989), as part of a viability study for bringing an alternative certification program to the University of Texas, El Paso, compared 63 Texas teachers who were traditionally certified, certified through an alternative certification program, or had an emergency permit. They found no differences between teachers on most outcomes. Likewise, a more rigorous study by Miller, McKenna, and McKenna (1998) used a match-paired design of 41 alternatively certified teachers and 41 traditionally certified teachers of fifth and sixth grade students, and found no differences in student achievement, teachers' instructional practices, or perceptions of ability between alternatively certified and traditionally certified teachers.

Two longitudinal studies (Raymond, Fletcher, & Luque, 2001; Decker, Mayer, & Glazerman, 2004) of alternatively certified teachers coming out of the TFA program found that TFA teachers and students fared as well as other certified teachers, new and veteran, on several measures. The first TFA study, conducted by Raymond, Fletcher, and Luque (2001), compared a group of TFA and non-TFA teachers in the Houston Independent School District. The researchers determined that the effects of being taught by a TFA teacher were generally positive and there was not a statistically significant difference in the achievement of students of non-TFA teachers and TFA teachers, as measured by the state standardized test (TAAS). In a more comprehensive study of TFA across six regions around the country and 17 high-poverty elementary schools, Decker, Mayer, and Glazerman (2004) compared elementary TFA teachers (n=44), in grades 1 through 5, to other elementary school teachers (n=56), new and veteran, in similar types of schools and determined through regression analysis that students of TFA teachers

made stronger gains in mathematics than did students of all other teachers and did as well as students of all other teachers in reading.

In summary, the evidence indicates that alternatively certified teachers receive favorable evaluations from supervisors and that alternatively certified teachers are as good as or even better than traditionally certified teachers in terms of effecting gains in student achievement. New research on TFA teachers suggests that alternatively certified TFA teachers are as better than fully certified teachers in mathematics, and as good as all other teachers, new and veteran, in general.

Overall, the evidence suggests that certification seems to identify good teachers in the area of high school mathematics. Emergency certification seems to identify poor-quality teachers. It seems that certification matters less for elementary school teachers. Though the research has found no clear difference in performance between alternatively certified teachers and those prepared through traditional routes, new analyses of TFA teachers suggest that these alternatively certified teachers perform better in mathematics and as good as all other teachers, regardless of certification status, in reading.

Teacher Certification and Instructional Practices

Very few empirical studies have been conducted that examine the relationship between teachers' certification status and their instructional practices. By definition, increased levels of certification indicate that teachers have undertaken further study, which makes available to them knowledge that they can operationalize through the types of classroom strategies they use. Of course, a teacher's practice is constrained by multiple factors, including control over curriculum decisions, students' needs, and available instructional resources (Cohen, Raudenbush, & Ball, 2003). The evidence

reviewed below focuses on the relationship between teachers' certification status and the classroom instructional practices they employ for young children.

In terms of certification status and the use of instructional practices for young children, Germino-Hausken, Walston, and Rathbun (2004), descriptive analysis of data from the Early Childhood Longitudinal Study (ECLS-K) found that teachers who did not hold either elementary or early childhood education certification used fewer constructivist types of instructional activities (such as activity centers) in their classrooms than did their certified kindergarten teacher counterparts. Even among certified teachers, important differences may remain in the actual implementation of practices promoted by programs or adopted by teachers (Vartuli, 1999). Vartuli (1999) examined various belief measures of kindergarten, first-, second-, and third-grade teachers and determined through correlational and regression analyses that teachers tended to support constructivist practices but may not implement them in their classrooms. She concluded that use of constructivist practices decreased as grade level increased.

Finally, Fidler (2002) used a hierarchical linear modeling approach to understand the relationship between teaching techniques and student achievement within the context of a class-size-reduction district reform. Using a sample of second and third grade students in the Los Angeles Unified School District, she determined that students of fully certified teachers had higher SAT-9 test scores and that certified teachers utilized teaching techniques (which she calls "individualization and engagement" strategies, drawn from a factor analysis of self-reported classroom activities) that deepened students' understanding in reading/language arts after controlling for language minority status, previous academic achievement, and SES. Additionally, she found that mathematics achievement was dependent mostly on prior achievement in mathematics.

In short, there is a paucity of studies related specifically to a teacher's type of certification and classroom instructional practices. The evidence reviewed above suggests that kindergarten teachers who did not hold early childhood or elementary certification used less constructivist-type instructional practices. Another study indicated that elementary teachers with full certification utilized individualization and engagement strategies that deepened students' understanding in reading. Some evidence suggests that even though constructivist practices are favored, teachers' actual implementation of them is uncertain.

Empirical Approaches for Measuring Instructional Practices

This section of the literature review focuses on empirical research related to the classroom instructional practices of public school teachers generally and early-grade teachers specifically. Though many interpretive studies have provided much-needed and nuanced information concerning teachers' classroom practices, the emphasis in this study is on understanding what large-scale surveys and secondary data analyses reveal about teachers' classroom practices. The studies examined in this section are divided into three groups. The first group focuses on the evidence from previous large-scale survey research studies that examine the effects of classroom practices on various student outcomes—most notably student achievement. The second group reviews several recent studies concerned with understanding how reform-oriented classroom practices are used by teachers and the relationship of these practices to student achievement. The third group of studies focuses specifically on the evidence related to the classroom practices of teachers of young students.

Before turning to the evidence, it is important to acknowledge the limitations of using empirical approaches and large-scale survey research to study teachers' classroom practices. First, studying teachers' classroom practices is quite difficult. Teachers' instructional practice is mediated by multiple factors, including, but not limited to, instructional policy (Mayer, 1999), teachers' control over the curriculum (Anderson, 1994), students' needs (Ellis & Worthington, 1994; Tomlinson, 2001), time allocated to instruction (Perie, Baker, & Bobbitt, 1997), and beliefs about teaching and learning (O'Loughlin, 1989; Woolley, Woolley, & Hosey, 1999). Second, these factors suggest a dynamic view of a teacher's classroom practice that is strongly influenced by factors that are not immediately observable and that are difficult to capture by the usual measurement instruments.

ECLS-K uses teacher self-report surveys of various classroom practices— instructional strategies, motivational techniques, and classroom activities. Although self-report data reveals the frequency with which teachers engage in various classroom practices, it does not capture the quality of practice. Burstein et al. (1995) found that “teachers' survey responses to fine-grained judgments of frequency of instructional practices were not entirely accurate” (p. 23). Though they are not frequently used in large-scale research studies, other classroom practice measurement instruments and strategies are used in quantitative research. Some of these instruments include participant observation and teacher logs—both written and electronic (Stecher et al., 2002)—and videotape studies (see Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). Despite these limitations, Mayer (1999) points out that the self-report surveys of teacher practices are invaluable for policymakers but suggests interpreting the data with caution.

Classroom Instructional Practices and Student Outcomes

The evidence reviewed below focuses on the use of empirical approaches to understand the relationship between teachers' classroom instructional practices and student outcomes—namely, student achievement. This section begins with an examination of early empirical studies and the contribution of school effectiveness research on the topic. Following this discussion, studies that used large-scale survey research to study the relationship between classroom instructional practices and student outcomes are reviewed. Although the focus of this study is on early elementary grade students, many previous studies using large-scale survey research have focused on secondary school students, and these studies are included as well.

The research tradition of examining classroom practices and various outcomes is relatively long. Medley and Mitzel (1959) investigated the instructional practices of 49 New York City first-year teachers to understand how various “teacher behaviors” correlate to student achievement. Similarly, Flanders (1960) studied the effects of “teacher influence” on students' attitudes toward learning and achievement. By today's standards, these early studies used relatively simple statistical procedures; the studies also were not longitudinal and did not control for several factors known to influence student achievement (e.g., SES, prior learning, and race).

Following these early studies, Coleman et al. (1966) and Jenks et al. (1972) began to use larger data sets to examine the effects of schools and teacher characteristics on student achievement. Both studies concluded that family and community factors, not schools, accounted for the greatest variance in student achievement. Controversy over these findings, coupled with unprecedented funding for research, spawned a body of

literature typically referred to as school effectiveness research, which revealed that schools and their teachers do in fact make small but significant differences in student outcomes (Edmonds, 1979; Brookover & Lezotte, 1979). Others (Brophy & Good, 1986; Purkey & Smith, 1983) have reviewed and synthesized this literature extensively. Although the school effectiveness research contributed significant knowledge about teachers' processes, much of the empirical evidence was based on relatively small samples of teachers and students at one point in time. Subsequent studies were able to provide nationally representative samples of students, teachers, and schools over time.

The specific evidence from large-scale survey research generally finds small statistically significant relationships between teachers' classroom practices (namely, the use of higher-order thinking strategies) and student achievement. Wenglinsky (2002) used a multilevel structural equation, in which SES and class size were used as controls, model of 1996 NAEP mathematics data for eighth grade students to determine that the effects of classroom practices were stronger than the effects of professional development. When classroom practices were added to other teacher characteristics, the effects were similar in magnitude to those of student background. Likewise, NCES (1996) researchers using NELS:88 data determined that 10th grade students of teachers who emphasized higher-order thinking skills performed better in mathematics but not in science.

Von Secker and Lissitz (1999) used NELS:90 data to study the effects of reform-oriented instructional practices (laboratory inquiry, critical thinking, and reducing teacher-centered instruction) of 10th grade high school science teachers on student achievement in science. Using a hierarchical linear modeling (HLM) approach, the researchers found that the use of reform-oriented instructional practices was not associated with significant differences in the mean achievement of a school's students.

Furthermore, the authors noted that as the frequency of reform-oriented practices increased, the disparity between the scores of minority and non-minority students and the scores of female and male students increased as well—raising serious equity concerns for the authors.

With respect to elementary school teachers, Henke, Chen, and Goldman (1999) used the Teacher Follow-Up Study 1994–1995 from the School and Staffing Survey in a descriptive study that examined several teacher characteristics (experience, degrees attained, and teachers' beliefs about student ability) and student characteristics (English proficiency, income, and race) on classroom practices. A particular emphasis was placed on teachers who engaged in reform-oriented practices as opposed to those who used conventional or traditional practices. Reform-oriented practices were defined as those emphasizing authentic assessment through portfolios, higher-order thinking skills, grouping strategies, and student-student talk. With respect to teacher characteristics, teachers who believed that their students were more able (i.e. gifted) used the recommended reform-oriented strategies less frequently than teachers whose students were not recognized as gifted. Teachers with more experience (more than 5 years) tended to use the recommended strategies more than teachers with less experience. Teachers with advanced degrees were more likely to use reform-oriented strategies than teachers with a four-year degree. Teachers who had participated in professional development on a reform-oriented strategy the previous year were more likely to use the recommended strategies.

In summary, the studies reveal that for high school students in mathematics, the use of higher-order thinking skills has a positive effect on 10th graders' mathematics

achievement. At the elementary level, at least one study indicates that teachers with more experience, who had a high percentage of minority students in their classrooms, and who had tentative beliefs about students' ability were more likely to engage in reform-oriented instructional practices. Interestingly, the use of reform-oriented practices at the high school level in mathematics seemed to create larger gaps in mathematics achievement between minority and non-minority students. The evidence pertaining to the use of reform-oriented strategies for elementary students is reviewed in greater depth in the next section.

Reform-Oriented Instructional Practices and Student Achievement

A key educational reform over the past decade has been the use of standards-based instruction (sometimes referred to as reform-oriented instruction). Advocates of reform-oriented instructional approaches argue that students will not be able to demonstrate proficiency and meet the required standards without innovations in teaching practice. The literature reviewed below tests this assumption.

Cohen and Hill (2000) determined that fourth grade teachers' instructional practices in mathematics are strongly influenced by instructional policy measures that shape instruction, curriculum, and assessments. Using an OLS regression model and student achievement data from the California Learning Assessment System (CLAS), the researchers determined that students of teachers who understood standards-based instruction, received professional development on it, and used instructional practices consistent with the reform (as opposed to conventional practices), scored higher on the CLAS. Similarly, McMillan (2003) utilized a convenience sample of 79 fifth-grade

teachers and 29 schools in a socially and economically diverse school system in Virginia. He administered an “instructional practices” survey in which teachers reported on a Likert-type scale their frequency of use of various classroom practices. He determined, primarily through a correlation analysis of identified instructional practice factors, students of teachers who engaged in reform-like practices (cooperative learning, formative assessments, essay tests, and direct instruction) in reading and mathematics scored higher on a high-stakes test of reading and mathematics.

In contrast to Cohen and Hill (2000) and McMillan (2003), a study by Klein et al. (2000) reported no observed difference in the frequencies of teachers’ use of reform or traditional practice. Klein et al. examined first-year findings from a study of reform-oriented curricular implementation in mathematics and science to determine whether there was a relationship between student achievement and teachers’ use of reform-oriented practices (cooperative learning groups, inquiry-based activities, use of materials and manipulatives, and open-ended assessment techniques). The researchers attributed the weak effects observed to methodological issues involved in large-scale research studies (Klein et al., 2000).

Like Klein et al. (2000), Berends, Chun, Schuyler, Stockly, and Briggs (2002) failed to find a relationship between reform-oriented practices and the reading and mathematics achievement of 3,800 fourth grade students. The researchers used a multilevel model to analyze the impact of reform-like instructional conditions—including teacher-reported collaboration, quality of professional development, and reform-like instructional practices—on fourth grade student achievement in reading and mathematics. Using controls for SES, prior achievement, and race, the findings indicated that pedagogical decisions were not related to increased achievement in mathematics or

reading on two measures of achievement (SAT-9 and the Texas Assessment of Academic Skills [TAAS]). The authors explained that the lack of a relationship may be due to several issues, chief among them the fact that the study was conducted when teachers were not familiar with the reform strategies.

In short, the use of reform-oriented strategies shows mixed effects on student achievement. The effects, although small, seem strongest with respect to high school students in the area of mathematics and weak for elementary school students in mathematics and reading. Although the researchers cite methodological challenges as a reason for the lack of an identified relationship, Gamoran, Porter, Smithson, and White (1997) put forth an alternative explanation which concludes that *what* is taught and tested is more important than *how* teachers teach it. Gamoran et al. (1997) systematically examined first year high school mathematics course content in an attempt to study opportunity-to-learn standards (OTL). OTL standards express a belief that student need to be taught what they are going to be tested on. The hope is that students will learn “past” the test and teachers will teach “past” what the test necessitates. They used a correlation analysis to determine how the association between topics covered on a mathematics achievement test and how OTL topics were taught by teachers to students. They determined that what ultimately matters is not how students are taught but that what they are taught as it relates to items covered on the test.

Reform-Oriented Instructional Practices Suggested for Teachers of Young Students

Like other professional organizations in this era of standards-based reform, early childhood educators’ organizations have also advocated for instructional practices (National Council of Mathematics Teachers [NCMT], 2005). The National Association

for the Education of Young Children (NAEYC) popularized the use of developmentally appropriate practices (DAP) in teaching in the early elementary grades (Bredekamp, 1987). This approach emphasizes constructivist child-centered approaches to teaching, learning, and behavior management (Lubeck, 1998; Bredekamp, 1993; Huffman & Speer, 2000).

Research concerning the effectiveness of constructivist pedagogy for young children seems mixed, depending on the outcome variables studied. Charlesworth's (1998) quasi-experimental study of primary teachers who used reforms that approximated DAP found that these teachers' students performed better on standardized tests than did students of teachers who used traditional instructional practices. With respect to gender differences and the use of constructivist curricula, Marcon (1999) examined three cohorts of inner-city pre-school children and three curricular approaches—academic directed, child initiated (i.e., constructivist), and a combination of the two. Marcon (1999) found that both boys and girls prospered when teachers used constructivist pedagogy, but boys seemed to benefit more from it.

In the affective domain, Burts et al. (1992) reported a lower frequency of stress behaviors and increased motivation of children in constructivist (i.e. DAP) classrooms. Burts et al. (1992) utilized a convenience sample of 200 teachers who self-reported the degree to which they used DAP. Observational data were collected from 12 classrooms which were classified as DAP or not based on teacher responses. Observer ratings were used to examine the validity of teacher responses concerning the characterizations of their classrooms and practices. The Classroom Child Stress Behavior Instrument (CCSBI) was used to measure child stress. The researchers determined, primarily through inferential analyses, that students in DAP classrooms reported less stress behaviors and

that Black students experienced more stress in non-DAP classrooms than White children. Within DAP classrooms there was no significant difference of stress behaviors between Black and White students.

Though the NAEYC's standards are used by accrediting bodies such as NCATE for the preparation of ECE teachers, some researchers have suggested that constructivist approaches have limitations, especially within an era of school accountability largely based on increasing students' test scores (Thompson, 2001). The instructional practices for raising tests scores usually focus on direct instruction rather than constructivist instruction (see Foorman, Fletcher, Francis, Schantschneider, & Mehta, 1998; Schrag, 1995). Furthermore, Diamond and Spillane's (2002) qualitative study involving participant observation and interview data of several Chicago public teachers and administrators revealed that teachers in high-poverty schools, where the consequences of poor test scores are more pronounced, often abandon constructivist-type practices by choice or necessity.

Pressley, Rankin, and Yokoi (1996) argued for a more balanced critique of instructional practices for elementary teachers, at least in literacy, rather than adopting one approach at the expense of the other. Pressley et al. surveyed 50 exemplary literacy instructors (based upon nominations by their supervisors) across the country and queried them about their classroom practices. Using factor analysis and several inferential procedures, the researchers identified patterns of practice in literacy instruction which revealed that while they espoused beliefs consistent with constructivism (i.e. whole-language reading approaches) they were more or less eclectic in their use of instructional practices and overwhelming used direct instruction.

In an effort to differentiate and quantify the relative effectiveness of various instructional practices, Marzano, Gaddy, and Dean (2000) conducted a meta-analysis in which they calculated effect sizes for a variety of instructional strategies. Their analysis yielded 10 groups of strategies, ranging from identifying similarities and differences to cooperative learning, each with an effect on student achievement. The strategies included both highly constructivist techniques such as cooperative learning as well as traditional techniques such as note taking. The researchers found a blend of practices most effective, with some strategies associated with traditional techniques producing higher percentile gains than nontraditional strategies.

In summary, some evidence indicates that teachers who use reform-oriented instructional approaches, such as developmentally appropriate practice, have students who score higher on standardized tests. However, other research indicates that students whose teachers use a blend of practices are likely to benefit the most.

CHAPTER III: METHODOLOGY

This study addresses three research questions related to teacher certification, instructional techniques, and student achievement:

1. How do the characteristics of first grade public school teachers, their schools, and their students vary by certification status?
2. Does a teacher's certification status predict the frequency of use of various classroom instructional activities in reading and mathematics?
3. Does a teacher's certification status predict how students will perform in reading or mathematics, or both?

The current chapter describes the methodology and statistical models used to address these questions. The chapter begins with a description of the Early Childhood Longitudinal Survey (ECLS-K) data set used in the analysis and provides information about the ECLS-K instruments that were a part of the survey. A description of the dependent, independent, and control variables used in the study follows. The chapter concludes with a discussion of the statistical procedures employed in the current analysis.

Description of Data

The data used in this dissertation are from ECLS-K, conducted by Westat, an independent research organization, for the U.S. Department of Education's National Center for Education Statistics (NCES). The purpose of the survey was to study the social and cognitive development of young children, as well as the effects of various education policies and practices on early development and learning. The study follows a

nationally representative sample of 22,782 kindergarteners in 1,277 kindergarten programs as they progress through the fifth grade. The study, which began in 1998–1999, includes data on these children, their families, their schools, and their teachers. Despite plans to follow the kindergarten cohort each year, NCES, due to fiscal and logistical constraints, restricted its surveys to the kindergarten, first grade (1999–2000), third grade (2001–2002), and fifth grade years (2003–2004).

The ECLS-K is based on a dual-frame, multistage sampling design (Denton & West, 2002). In the initial stage, developers selected 100 primary sampling units (PSUs) from a national sample of PSUs from which schools and students within those schools were randomly drawn. Private schools were identified using the Private School Survey and public schools using the Common Core Data Survey. An average of 23 kindergarteners was selected in each of the originally sampled schools. Oversampling of Asian children and private kindergartens was performed to permit generalizability to smaller populations of children and programs. During the first follow-up, NCES collected data on all kindergarteners who attended first grade in sampled schools and a random sample of roughly half the students who transferred to new schools. To compensate for the loss of students in the original sample, NCES “freshened” the first grade sample (i.e. drew additional random samples of children from the same schools) to enhance internal and external validity (Denton & West, 2002).

Public release data are available for five collection periods—fall kindergarten, spring kindergarten, fall first grade, spring first grade, and spring third grade. Spring fifth grade data are currently being collected. Response rates during the base-year period were generally good. Denton and West (2002) report that the school response rate was 74.2%, the student response rate 92%, the parent rate 89%, and the teacher rate 82%. This

dissertation uses longitudinal data gathered from the spring kindergarten (1999) and spring first grade (2000) survey administrations.

Large-scale survey research, such as the ECLS-K, is subject to both sampling and nonsampling errors. Nonsampling errors result largely from measurement error or human error in data collection and processing. Sampling errors occur from nonresponse and inefficiencies in the original sampling design. In order to produce accurate national estimates from the sample, as well as to compensate for the oversampling of small populations, the ECLS-K data set includes a series of statistical weights to use in analyses. According to Denton and West, “Weighting the data adjusts for unequal selection probabilities at the school and child levels and then adjusts for school, child, teacher, and parent non-response” (2002, p. 28). This study used precalculated child-level longitudinal sample weights computed for use with spring kindergarten and spring first grade data.

Instruments

A range of ECLS-K instruments were used in this study. They included the ECLS-K Direct Child Assessment administered in the spring of the children’s kindergarten and first grade years, the First Grade Teacher Questionnaire, and the school-level Administrator Questionnaire. The ECLS-K Direct Child Assessment includes measures of students’ proficiency in mathematical thinking, reading (language and literacy), and general knowledge (science and social studies). The First Grade Teacher Questionnaire includes teacher self-report data about certification type, training, instructional approaches, and other demographic data, and the Administrator

Questionnaire includes basic information about school organization, enrollments, and policies. Explanations of each instrument follow below.

Direct Child Assessment

The ECLS-K Direct Child Assessment contains three batteries of content-area achievement tests, focusing on reading, mathematics, and general knowledge. Each battery is administered through a single one-on-one hour-long testing period using computer-assisted interviews (CAI). The assessment follows a two-stage design in which the child answers a series of routing questions in the first stage to determine the appropriate level of difficulty for the questions in the second stage. Children responded to the questions by pointing to answers on a small easel and were not asked to write or explain anything. Paper and pencil and other manipulatives were provided for portions of the mathematics battery.

Prior to the content batteries, students were administered a language screener, the Oral Language Development Scale (OLDS), to ensure that they could understand and respond to items in English. If the primary language of the home was not English, the child was given the English OLDS. Students who met the cut score for the English OLDS were then given the content batteries in English. Students who did not reach the cut score in English and spoke Spanish were given a version of the Spanish OLDS to determine their native language proficiency, and were administered the mathematics battery in Spanish. Students with cognitive disabilities (i.e. special education students) received accommodations or alternative assessments consistent with the documented disability (NCES, 2002).

The reading and mathematics achievement batteries provided a range of IRT scale scores that were used as the primary measures of student achievement in the study. These IRT scale scores were used to estimate general achievement gains in reading and mathematics between different assessment periods.

Reading battery. The reading battery included test items to measure basic skills (print familiarity, letter and word recognition, beginning and ending sounds, rhyming sounds), vocabulary (receptive vocabulary), and comprehension (listening, words in context). The bulk of questions focused on comprehension (50%), followed by basic skills (40%) and vocabulary (10%) (NCES, 2002).

The ECLS-K data set includes an IRT-estimated overall score in reading for children who completed the battery, as well as an estimate of students' proficiency in five areas of reading skills: (a) letter recognition (identifying upper- and lowercase letters by name), (b) beginning sounds (associating letters with sounds at the beginning of words), (c) ending sounds (associating letters with sounds at the end of words), (d) sight words (recognizing common words by sight), and (e) comprehension of words in context (reading words in context).

Mathematics battery. The mathematics battery included (a) test items to measure conceptual knowledge (what something is), (b) procedural knowledge (how to perform an operation), and (c) problem solving (applying both conceptual and procedural knowledge). A majority of the mathematics test items involved questions on number sense, number properties, and operations (50%), followed by measurement, geometry, and spatial sense (35%), and data analysis, statistics, probability, patterns, algebra, and functions (15%).

Along with the IRT estimate of students' overall scores in mathematics, the ECLS-K Direct Child Assessment also addresses proficiency levels in five areas: (a) number and shape (identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting of up to ten objects), (b) relative size (reading all single-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare objects), (c) ordinality and sequence (reading two-digit numerals, recognizing the next number in a sequence, identifying the ordinal position of an object, and solving a simple word problem), (d) addition/subtraction (solving simple addition and subtraction problems), and (e) multiplication/division (solving simple multiplication and division problems and recognizing more complex number patterns). As with the reading battery, proficiency at higher levels implies proficiency at lower levels in mathematics.

Spring First Grade Teacher Questionnaire

Each teacher of a sampled child received a three-part self-administered questionnaire. The first questionnaire, Part A, queried teachers about general classroom characteristics such as the composition and demographics of the classroom. Part B, utilized most heavily in the study, collected background data relating to the teacher's training, certification status, perceptions of school culture and climate, and typical classroom activities (i.e. teaching strategies and the time spent teaching certain skills and knowledge). Part C asked teachers to rate the academic performance of the sampled children in their class in the three content areas assessed as part of the study.

Spring Administrator Questionnaire

The Spring Administrator Questionnaire collected information about school-level data such as school composition, policies, and practices. Pertinent items of interest from this questionnaire include (a) school characteristics, (b) teacher/staff characteristics, and (c) uses of instructional grouping and individualization.

Variables

This section discusses the dependent, independent, and control variables used in the current study.

Dependent Variables

Table 3.1 presents a summary of the primary dependent variables employed in the study—the first grade IRT scores for mathematics (C4R2MCSL) and reading achievement (C4R2RCSL). These variables provide an estimate of children’s knowledge and skills in reading and mathematics at the end of the first grade, measured by the direct student assessment. In the analysis, these variables have been standardized and expressed in a z-score metric with means of 0 and standard deviations of 1.

Table 3.1. Description and Summary of Overall Achievement Variables^a

ECLS-K variable label	Description	Unstandardized values ^b	Standardized values
Mathematics Achievement, C4R2MCSL	Spring 1st grade mathematics IRT scale score	Mean = 43.2, SD = 7.7 Range = 10.7–60.4	Mean = 0, SD = 1
Reading Achievement, C4R2RCSL	Spring 1st grade reading IRT scale score	Mean = 55.3, SD = 11.6 Range = 16.0–88.1	Mean = 0, SD = 1

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user’s file).

^b Based on analytic sample (n = 3,151 teachers). Scores are averaged across each teacher’s students.

Independent Variable

A first grade public school teacher’s certification status (B4TYPCER) was the primary independent variable used in the study. Teachers in the ECLS-K study were asked to identify their current level of certification from a list of five choices: (a) none; (b) temporary, probationary, or emergency; (c) completion of an alternative certification program; (d) regular or standard state certificate; and (e) advanced professional certificate. These certification levels, as shown in Table 3.2, are recoded into a new variable (CERTIFIC) with three new categories: (a) emergency—“none,” “temporary,” and “emergency” categories;³ (b) standard certification—“regular” or “standard” state certification; and (c) advanced certification—“advanced professional.” Teachers who

³ The certification categories developed ECLS-K is problematic in this study and noted again in the limitations section. In short, for many credentialing authorities probationary teachers have satisfied the usual requirements for initial certification (see Table 1.1) and need only satisfactory classroom experience for the next level of certification. Likewise, the temporary certification designation can refer also to previously fully-certified teachers who have transferred into a system from another one and may have met all but a few certification requirements for their new placements. Therefore, readers should exercise caution in interpreting the results of emergency certification category used in this study.

reported that they had completed some form of alternative certification were not included, because too few ($n = 21$) of them existed in the analytic sample.

Table 3.2. Description of Independent Variable^a

ECLS-K variable/recoded variable label	Description	Values ^b
Recoded Certification Status, CERTIFIC	Recoded to three certification categories: 1 = Emergency Certification (which includes none, temporary, emergency, provisional); 2 = Standard Certification (includes regular/standard); 3 = Advanced Certification (includes advanced professional)	Emergency Cert, $n = 251$ Standard Cert, $n = 2,644$ Advanced Cert, $n = 257$

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user’s file).

^b Based on analytic sample ($n = 3,151$ teachers).

Control Variables

There were three groups of control variables. These groups included (a) student characteristics, (b) teacher characteristics, and (c) classroom and school characteristics. Each group is explained in greater detail below.

Student characteristics. The variables for student characteristics present critical demographic data and previous student achievement data widely believed to be associated with student achievement. These data are reported in the administrator and teacher questionnaires and are summarized in Table 3.3. In the current study, race was recoded from the ECLS-K–provided race composite variable, RACE, into a new variable, RRACE, which represents the proportion of students in each teacher’s classroom who

were nonwhite. Similarly, students' gender (GENDER) was recoded as the proportion of students in the teachers' classrooms who were male. Student socioeconomic status was measured using a continuous SES measure provided by ECLS-K, which again was averaged for the students in each teacher's classroom. Achievement scores in reading and mathematics from the prior year (kindergarten) were assessed in the same manner as the primary dependent variables, which were averaged for the students in each teacher's classroom. Since the spring first grade ECLS-K Direct Child Assessment was administered over a relatively long window (from March to June), a measure of elapsed time (ELAPSE1) was created by using the assessment date variables to calculate the number of days between the two direct assessments. ELAPSE1, which was averaged for the students in each teacher's classroom, helps control for possible differences in opportunities to learn between students tested at different time points.

Teacher characteristics. The variables associated with teacher characteristics provide both personal demographic data and professional information about teachers (see Table 3.4). These characteristics were used to control for possible contributing effects of certification status on achievement and teacher practices. These teacher-related variables are self-report data from Teacher Questionnaire A. The teacher characteristics included race (TRACE), which was recoded to express the proportion of nonwhite teachers in the sample; age (B4AGE); education level (S4EDLVL), which was recoded as a teacher's highest degree (TEACHMAS)—less than a master's degree versus more than a master's degree; and number of years teaching (B4YRSTC), which was recoded to express the following three categories: less than 2 years, between 2 and 6 years, and more than 6 years.

Table 3.3. Description and Summary of Student Characteristic and Achievement Score Variables^a

ECLS-K variable/ recoded variable label	Description	Unstandardized values ^b	Standardized values or frequencies
Student Race Recoded, RRACE	Proportion of nonwhite students	NA	Nonwhite, <i>n</i> = 1,258 White, <i>n</i> = 1,893
Student Gender Recoded, GENDER	Proportion of male students	NA	Male, <i>n</i> = 1,607 Female, <i>n</i> = 1,544
Student SES W1SESL	Continuous SES measure	Mean = 0, SD = .6	Mean = 0, SD = 1
Spring Kindergarten Mathematics Achievement, C3R2MCSL	Spring kindergarten mathematics IRT scale score	Mean = 27.1, SD = 7.2 Range = 8.2–58.1	Mean = 0, SD = 1
Spring Kindergarten Reading Achievement, C3R2RCSL	Spring kindergarten reading IRT scale score	Mean = 32.1, SD = 8.6 Range = 12.5–82.2	Mean = 0, SD = 1
Time Lapse Between Kindergarten and 1st Grade Assessments, ELAPSE1	Computed variable; time, in days, between administration of kindergarten and 1st grade assessments	Mean = 367, SD = 21.2 Range = 294–437	Mean = 0, SD = 1

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user’s file).

^b Based on analytic sample (*n* = 3,151 teachers)

Table 3.4. Summary of Teacher Characteristic Control Variables^a

Variable label	Description	Values ^b
Teacher's Race, TRACE	Self-reported race of teacher; recoded to 0 = white, 1 = non-white	White, <i>n</i> = 2,685 Nonwhite, <i>n</i> = 466
Teacher's Age, B4AGE	Teacher's self-reported age	Mean = 41.07, SD = 10.78 Range = 23.00–65.00
Highest Degree, TEACHMAS	Proportion of teachers with a master's or beyond	< Master's, <i>n</i> = 1,926 Beyond master's, <i>n</i> = 1,225
Total Years Teaching Experience, B4YRSTC	First grade teachers' self-reports of total years teaching; recoded to 1 = less than 2 years, 2 = between 2 and 6 years, 3 = more than 6 years	<2 yrs., <i>n</i> = 377 2–6 yrs., <i>n</i> = 2,208 >6 yrs., <i>n</i> = 566

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user's file).

^b Based on analytic sample (*n* = 3,151 teachers)

Instructional practice variables and school and classroom characteristics. Since the First Grade Teacher Questionnaire presents more than 25 instructional activities and techniques teachers, a factor analysis was conducted to create a few instructional technique composite variables to use in the analysis. At the kindergarten level, Germino-Hausken, Walston, and Rathbun (2004) have successfully used similar items from the ECLS-K Direct Child Assessment to create factors that tap different types of teacher practices, such as an emphasis on comprehension skills, phonics, student-centered instruction, and mixed achievement grouping. Similarly, in mathematics, they created factors that emphasize numbers and geometry, traditional practices and computation, measurement and advanced topics, and student-centered instruction (Hamilton &

Guarino, 2004). Fidler (2002) used a similar approach with different data and found similar instructional practice variables. Her analysis yielded a set of instructional practice factors across reading and mathematics, which included (a) individualization and engagement; (b) redundancy, practice, and modeling; and (c) classroom management. Results from the factor analysis, provided in Chapter 4, indicated that three mathematics instructional factors and three reading instructional factors were necessary and sufficient to capture teacher differences in instructional activities.

Since the instructional techniques teachers use and student achievement seem influenced by the context of the classroom, key classroom-level variables were included in the analysis. These variables attempt to provide a broad description of ethnic and socioeconomic diversity, and class size. The variables were primarily collected from teacher-reported data in Teacher Questionnaire A, the Administrator Questionnaire, and ECLS-K Direct Child Assessment imputed values. Table 3.5 presents a summary of these variables, which include class size (CLASSIZ), percentage of minority students in the classroom (PERCMINO), and whether or not the school received Title I funds (TITLE1).

Table 3.5. Summary of School and Classroom Characteristic Control Variables^a

Variable name	Description	Values ^b
Class Size, CLASSSIZ	Total class enrollment	Mean = 20.4, SD = 3.6
Title 1 Funds, TITLE1	School receives Title I funds 0 = school did not receive Title I funds; 1 = school received Title I funds	Did not receive, $n = 805$ Received, $n = 1,893$
Percentage of Minority Students, PERCMINO	Percent minority students 1 = 0% to 50%, 2 = 51% to 100%	>50%, $n = 1,961$ <50%, $n = 1,190$

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user’s file).

^b Based on analytic sample ($n = 3,151$ teachers)

Analytic Sample

The students included in the analysis had kindergarten experience in 1998 and were promoted to first grade in a public school in the fall of 1999. Additionally, these students had first grade spring Direct Child Assessment scores as well as corresponding teacher questionnaire data. Teachers included in the analysis reported that they fit in one of the ECLS-K Direct Child Assessment certification categories. To provide for accurate comparisons, the sample includes only public school students who were administered the reading assessment in English in both the spring kindergarten and spring first grade administrations. The sample also excludes Spanish-speaking students’ mathematics scores.⁴

⁴ Students who did not pass the OLDS were not given the opportunity to take the reading battery. Spanish-speaking students, however, were administered the mathematics battery in Spanish. The batteries were not available in languages other than Spanish and English.

Statistical Procedures

A variety of statistical procedures were employed in the current study. Descriptive statistics were provided for all study variables. In addition, statistical significance tests (correlations, chi-squares, and an analysis of variance between groups [ANOVA]) were conducted between teacher certification and every other study variables as a preliminary step in determining the relationships between teacher certification status and the other study variables. Then, factor analyses were conducted separately for the mathematics instructional practices and the reading instructional practices. The primary inferential tool employed to address the research questions of the current study was OLS regression, which is a common approach used in social science research. It allows the researcher to understand the predictive strength of the independent variable (certification status) on the dependent variable (student achievement), while at the same time incorporating the effects of other relevant variables (controls).

CHAPTER IV: FINDINGS

This chapter presents the findings generated by the analytic procedures discussed in Chapter 3. To review, this study is concerned with the following research questions:

1. How do the characteristics of first grade public school teachers, their schools, and their students vary by certification status?
2. Does a teacher's certification status predict the frequency of use of various classroom instructional activities in reading and mathematics?
3. Does a teacher's certification status predict how students will perform in reading or mathematics, or both?

The current chapter addresses these questions by first examining the characteristics of the students, classrooms, and schools of teachers with different certification statuses, as well as the characteristics of the teachers themselves. Second, the chapter focuses on the principal component analyses (PCA) used to identify factors related to teachers' instructional practices in mathematics and in reading. These analyses provide a basis from which to examine the second research question—namely, the extent to which classroom instructional activities vary by certification status. Finally, the chapter ends with a discussion of the findings from regression analyses involving the relationships between the student, teacher, and school variables and mathematics and reading achievement.

Descriptive Analyses of Students, Teachers, Schools, and Classrooms

Student Descriptives

Tables 4.1 through 4.3 provide descriptive statistics for students, teachers, classrooms and schools by teachers' certification status. Descriptive statistics for the students are provided in Table 4.1. Rows compare the descriptive statistics of students of teachers with three certification types—emergency, standard, and advanced. In order to determine whether there is a relationship between certification status and student characteristics, two different statistical tests are used: chi-square in the case of categorical variables, and one-way ANOVA with post hoc contrasts in the case of continuous variables (emergency vs. standard and advanced vs. standard). The results of these tests are indicated in the table.

When examined at the teacher level ($n = 3,151$ teachers), first grade students are far more likely to be taught by teachers with standard certification status (82.2%) than by teachers with either emergency (8.7%) or advanced (9.1%) certification status. However, specific populations of students are more likely than other students to be taught by either emergency or advanced certification status teachers. Considering the demographic variables first, male students are slightly more likely to have teachers with emergency (8.3% vs. 8.0%) and advanced (8.8% vs. 8.1%) certification status than female students, nonwhite students are more likely than white students to have teachers with emergency certification status (10.4% vs. 8.0%), and white students are more likely than nonwhite students to have teachers with advanced certification status (8.4% vs. 7.2%).

Table 4.1. Background Characteristics of First Grade Students in the Analytic Sample (N = 3,151)^a

Student characteristics	Emergency certified teachers (8.7%)	Standard certified teachers (82.2%)	Advanced certified teachers (9.1%)
Gender of Student, GENDER			
Male	8.3%	82.9%	8.8%
Female	8.0%	83.9%	8.1%
Race of Student, RACE^{b***}			
White	8.0%	83.9%	8.4%
Nonwhite	10.4%	82.4%	7.2%
Socioeconomic Status, ZSES^{c*}			
	-.15*	.01	.08*
Assessment Lapse (Days Between Kindergarten and 1st Grade), ELAPSE			
	366	365	364
Kindergarten Mathematics Achievement (z), M_IRT_2Z^{c***}			
	-.27***	.03	-.07*
First Grade Mathematics Achievement (z), M_IRT_4Z^{c***}			
	-.18***	.02	-.06*
Kindergarten Reading Achievement (z), R_IRT_2Z^{c***}			
	-.23***	.03	-.09**
First Grade Reading Achievement (z), R_IRT_4Z^{c***}			
	-.25***	.23	-.02**

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user's file).

^b Significance based on chi-square test with 2 degrees of freedom (GENDER and RACE).

^c Significance based on one-way between groups ANOVA (Emergency vs. Standard; Standard vs. Advanced) with Tukey HSD post hoc tests of mean differences.

* $p < .05$, ** $p < .01$, *** $p < .001$

There is also a statistically significant difference in the mean socioeconomic status (SES) of students taught by teachers with standard compared to advanced certification status. Because SES is expressed in a z-score metric, mean differences can be interpreted as a percentage of a standard deviation (SD) in the distribution of SES across all students.⁵ As Table 4.1 indicates, the mean SES of students being taught by advanced certification status teachers is higher than the mean SES of students being taught by teachers with standard certification status. The one-way ANOVA for mean student SES by teacher certification status is significant ($p = .04$). Tukey HSD post hoc tests for the difference between the mean SES of students with standard versus advanced and emergency versus standard certification status teachers were significant at the .05 level.

Table 4.1 displays even larger differences in the achievement characteristics of students taught by teachers with different certification statuses. These differences range from roughly one-third to one-half of an SD, indicating that *both* emergency and advanced certification status teachers are more likely to have lower-achieving students in their classrooms. Although there is no relationship between teachers' certification status and the amount of time that elapsed (in days) between the spring kindergarten assessment and the spring first grade assessment, there is an overall pattern of teachers with emergency or advanced certification status teaching students with lower end-of-kindergarten and lower end-of-first-grade achievement scores. Based on the results of the one-way ANOVA with contrasts, the mean achievement scores of students differed by teacher certification status ($p < .001$) for both reading and mathematics, with the mean

⁵ All achievement measures are expressed similarly in this z-score metric.

scores of students of teachers with emergency certification status ($p < .001$) and advanced certification status ($p < .05$) being significantly lower than the mean scores of students of teachers with standard certification status. Emergency certified teachers taught students with the lowest end-of-kindergarten mathematics scores (-.27 SD vs. .03 SD), the lowest end-of-kindergarten reading scores (-.23 SD vs. .03 SD), and the lowest end-of-first grade mathematics scores (-.18 SD vs. .02 SD). Additionally, teachers with emergency certification status taught students with the lowest end-of-first-grade reading scores (-.25 SD vs. .23 SD).

Teacher Descriptives

Table 4.2 summarizes the demographic and educational characteristics of teachers in the analytic sample. Each row provides descriptive statistics for teachers classified by their certification status. As in the previous table, chi-square analyses in the case of categorical variables and one-way ANOVA with post hoc contrasts in the case of continuous variables (emergency vs. standard and advanced vs. standard) are used to test whether there is a statistical relationship between these teacher characteristics and certification status. Note that these analyses are presented at the teacher level ($n = 3,151$ teachers).

Table 4.2. Demographic and Education Characteristics of First Grade Teachers in the Analytic Sample ($N = 3,151$)^a

Teacher characteristics	Emergency certified teachers (8.7%)	Standard certified teachers (82.2%)	Advanced certified teachers (9.1%)
Race of Teacher, TEACHRACE ^{b***}			
Nonwhite	15.1%	76.3%	8.6%
White	6.7%	85.2%	8.0%
Mean Teacher's Age ($n = 3,038$), TEACHAGE ^{c***}	33 ^{***}	41 ^{***}	44 ^{***}
Teacher's Degrees Earned, TEACHMAS ^{c***}			
Less Than Master's Degree	10.2%	87.0%	2.8%
Master's Degree or Higher	4.4%	79.0%	16.6%
Teaching Experience ($n = 3,132$), YRSTEACH ^{c***}			
Less Than 2 Years	32.4%	65.4%	2.2%
Between 2 and 6 Years	16.0%	78.8%	5.2%
More Than 6 Years	2.5%	87.6%	9.9%

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K-1) (public-user's file). *Note.* Sample sizes are provided for variables with less than complete data.

^b Significance based on chi-square test with 2 degrees of freedom (TEACHRACE, TEACHMAS) and 4 degrees of freedom (YRSTEACH).

^c Significance based on one-way between groups ANOVA (Emergency vs. Standard; Standard vs. Advanced) with Tukey HSD post hoc tests of mean differences.

* $p < .05$, ** $p < .01$, *** $p < .001$

Although a majority (85.2%) of the teachers in the sample is white, significant differences do emerge when the teachers are grouped by certification level. Emergency certified teachers are more likely to be nonwhite (15.1% vs. 6.7%), while teachers with advanced certification are slightly more likely to be white (8.6% vs. 8.0%). The mean

teacher's age in the sample is approximately 41 years. Not unsurprisingly, emergency certified teachers tend to be younger, with a mean age of 33 years, while teachers with advanced certification tend to be older, with a mean age of 44 years. The one-way ANOVA indicates that there was a statistically significant difference in the age of the teachers in the three certification groups ($p < .001$). Tukey HSD tests indicate that emergency certified teachers are significantly younger than standard certified teachers ($p < .001$), and teachers with advanced certification are significantly older than teachers with standard certification ($p < .001$).

Table 4.2 also indicates noticeable differences between teachers with different certification statuses and their qualifications—again, a difference that might be anticipated given general perceptions about the relationship between certification status and qualifications. Overall, 38.9% of the teachers in the sample possess an advanced degree (master's degree or higher). However, emergency certification status teachers are significantly less likely to have an advanced degree than a lower degree (10.2% vs. 4.4%), while, conversely, teachers with advanced certification are more likely to have an advanced degree than a lower degree (16.6% vs. 2.8%). The test of independence between teachers' certification level and educational attainment is statistically significant ($p < .001$). Teachers' levels of experience—categorized as less than 2 years, between 2 years and 6 years, and more than 6 years—also vary systematically by certification status, with 32.4% of teachers with less than 2 years of experience being emergency certification teachers compared to 2.2% being advanced certification status teachers. The observed percentages of emergency, standard, and advanced certification status teachers grouped by experience were statistically significant ($p < .001$).

Classroom and School Descriptives

Descriptive statistics for teachers' classrooms and schools are provided in Table 4.3. Each row compares the descriptive statistics of various aspects of classrooms and schools by teachers' certification status. Chi-square analyses in the case of categorical variables and one-way ANOVA with post hoc contrasts in the case of continuous variables (emergency vs. standard and advanced vs. standard) are used to test whether there is a statistical relationship between the characteristics of classrooms and schools and the certification status of teachers. As with the previous tables, these analyses are presented at the teacher level.

The average class size for the sample is 20 students per classroom, and the sample class size mean is consistent across all certification types—that is, class size did not vary systematically with the certification status of teachers. Emergency certification teachers, however, are more likely to teach at schools with higher minority enrollments (11.6% vs. 5.8%) and schools receiving Title I funds (8.3% vs. 5.6%), whereas advanced certification teachers are less likely to teach in schools with high minority enrollments (6.3% vs. 9.3%) or in schools that receive Title I funds (7.6% vs. 9.3%). The tests of independence between teachers' certification status and the school's minority enrollment level ($p < .01$) and receipt of Title I funds ($p < .05$) are statistically significant.

Table 4.3. Characteristics of First Grade Teachers' Classrooms and Schools in the Analytic Sample (N = 3,151)^a

Teachers' classroom and school characteristics	Emergency certified teachers (8.7%)	Standard certified teachers (82.2%)	Advanced certified teachers (9.1%)
Average Class Size, CLASSIZ	21	20	20
Minority Composition of School (<i>n</i> = 3,118), PERCMINO ^{b***}	5.8%	85.0%	9.3%
0 to Less Than 50%	11.6%	82.1%	6.3%
50% or More			
Received Title I Funds (<i>n</i> = 2,697) TITLE1 ^{b*}			
Yes	8.3%	84.1%	7.6%
No	5.6%	85.1%	9.3%

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K-1) (public-user's file). *Note.* Sample sizes are provided for variables with less than complete data.

^b Significance based on chi-square test with 2 degrees of freedom.

* $p < .05$, ** $p < .01$, *** $p < .001$

Summary

In sum, the descriptive analyses reported above highlight important distinctions between teachers with emergency, standard, and advanced certification. Some of the differences were anticipated, while others were not. In terms of student demographic characteristics, the finding that emergency certified teachers are more likely to teach lower-SES students, students of color, and lower-achieving students is consistent with evidence presented earlier (see Qu & Becker, 2003; Darling-Hammond et al., 2001). Advanced certification status teachers, in contrast, are more likely to teach socially and

economically advantaged students. Intuitively, there is reason to suspect that teachers with advanced certification would have the highest-performing students, but these data do not support such a belief. The students of teachers with advanced certification, when compared to the students of teachers with standard certification, were more likely to score below the mean in both first grade and kindergarten mathematics and reading achievement tests, though not as far below the mean as were students taught by emergency certified teachers. The findings regarding the characteristics of teachers with different certification statuses were expected, with emergency certified teachers being younger, being more likely to teach in schools with high minority enrollments, and having less experience and fewer degrees than teachers with standard or advanced certification.

Principal Component Analyses of Identified Instructional Practices

The principal component analyses (PCA) presented in Tables 4.4 through 4.8 identified a variety of instructional practices in reading and mathematics that ranged from traditional practices such as arithmetic and phonics to higher-level, student-centered practices that require students to construct meaning and use knowledge meaningfully. Each analysis was conducted by randomly splitting the student sample ($n = 10,170$) into two samples—Subsample 1 ($n = 5,120$) and Subsample 2 ($n = 5,050$). The initial analysis was performed on Subsample 1, in which one-, two-, three-, four-, and five-component solutions were examined for both the reading and mathematics items. Varimax rotations were used on all multi-component solutions. Based on scree plots and interpretability, the three-component solution was selected for both reading and mathematics. Items that loaded less than .40 on all three components, or that loaded greater than .40 on two or

more components, were excluded and the analysis was rerun on Subsample 1. This process was repeated until a stable solution was achieved, at which point the analysis was run on Subsample 2 to examine replicability. Then the two subsamples were combined and the final analyses were conducted on the entire sample.

Reading Instructional Practices

The three reading instructional practice factors generated by the reading PCA describe a range of classroom activities that include both traditional and reform-oriented instructional practices. Of the original 48 reading activities provided by ECLS-K, 18 were subsequently removed, leaving 30 items with loadings greater than .40 on a single component. Table 4.4 shows the results of the 30-item analysis on Sample 1 and Sample 2. As can be seen, the solutions for the two samples were very similar in terms of the percentage of variance explained by each component and the internal consistency reliability coefficients (Cronbach's alpha).

Table 4.5 presents the component loadings and other statistics for the three-component solution for the reading activities. The three components also explained very similar percentages of variance and had very similar internal consistency reliabilities in each subsample. The first component, which includes many reform-oriented practices, was labeled *Student-Centered Reading Instruction*. This factor explains 13.38% of the variance among the reading activity items and had an internal consistency reliability of .80. The second component, labeled *Fluency and Comprehension Instruction*, which

Table 4.4. Preliminary Principal Component Analysis Results for Reading Instruction

	Component 1	Component 2	Component 3	Total
<u>Sample 1 (n = 5,120)</u>				
Sum of squared loadings	3.95	3.71	3.22	10.88
Percentage of variance explained	13.2%	12.4%	10.7%	36.3%
Internal consistency reliability	0.79	0.77	0.82	—
Number of items	11	13	6	32
<u>Sample 2 (n = 5,050)</u>				
Sum of squared loadings	4.02	3.66	3.31	10.98
Percentage of variance explained	13.4%	12.2%	11.0%	36.6%
Internal consistency reliability	0.80	0.77	0.83	—
Number of items	11	13	6	30

Note. Internal consistency reliability is for an equally weighted composite of all variables with loadings greater than or equal to .40 in the analysis of Sample 1.

focuses on intermediate beginning reading skills, explains 12.19% of the variance and has an internal consistency reliability coefficient of .77. The third component, labeled *Letter Sense Instruction*, includes foundational reading skills and explains 10.89% of the variance with an internal consistency reliability of 82.

Table 4.5. Final Principal Component Analysis Results for Reading Instruction

Instructional practices	Components		
	Student-centered literacy instruction	Fluency and comprehension instruction	Letter sense instruction
Write stories/report	.71	.07	-.06
Publish own writing	.66	.03	-.01
Work related to books	.63	.11	.01
Projects in small groups	.61	.06	.09
Story has a beginning, middle, and end	.60	.22	-.05
Long projects	.55	-.04	.11
Write in journal	.54	.04	.01
Perform play/skits	.53	.03	.09
Write with invented spellings	.52	.11	.02
Mixed-level groups	.49	.06	.07
Retell stories	.44	.34	.06
Read phonetic patterns	.05	.64	.14
Vocabulary	.03	.57	.03
Read controlled vocabulary	-.03	.57	.08
Reading aloud fluently	.06	.56	-.04
Read patterned text	.24	.54	.16
Read aloud	.09	.51	.03
New vocabulary	.10	.50	.14
Conventional spelling	.04	.49	.00
Work on phonics	-.08	.45	.21
Identify main idea of story	.34	.45	.01
Write from dictation	.03	.44	.08
Use cues for comprehension	.32	.43	.00
Alphabetizing	.16	.42	.05
Alphabet and letter recognition	.07	.05	.86
Matching letters to sounds	.04	.06	.74
Writing own name	.04	.04	.74
Convention of print	.15	.04	.71
Work on letter names	.02	.20	.68
Writing alphabet	-.01	.22	.55
Sum of squared loadings	4.01	3.66	3.27
Percentage of variance explained	13.38	12.19	10.80
Internal consistency reliability	.80	.77	.82
Number of items	11	13	6

Note. Internal consistency reliability is for an equally weighted composite of all variables with loadings greater than or equal to .40 on each particular component.

Mathematics Instructional Practices

Tables 4.6 and 4.7 describe the results of the PCA used to generate the mathematics instructional practice factors, including the factor loadings for each instructional practice. Of the original 48 mathematics activities provided by ECLS-K, 16 of these items were subsequently removed, leaving 32 items with loadings greater than .40, with one exception—the “frequency of mixed operations” item. Ultimately, this item was dropped because it loaded only .38 on its intended component for the entire sample. Table 4.6 presents a summary of the three-component analysis of the mathematics activity items in the two samples. As can be seen, the solutions for the two samples are very similar in terms of the percentage of variance explained by each component and the internal consistency reliability coefficients. The samples were combined and the three-component model was recomputed.

Table 4.7 presents the loadings and summary statistics for the final three-component solution of the 31 mathematics items. The first component, labeled *Student-Centered Mathematics Instruction*, includes reform-oriented and constructivist-type activities. This factor explains 12.28% of the variance and had an internal consistency reliability coefficient of .79. The second mathematics component, labeled *Computation Instruction*, includes arithmetic skills and activities. This component explains 9.80% of the variance and had an internal consistency reliability coefficient of .72. The third component, labeled *Number Sense Instruction* consists primarily of foundational mathematics skills and activities; it explains 15.28% of the variance among the mathematics items and has an internal consistency reliability coefficient of .84.

Table 4.6. Preliminary Principal Component Analysis Results for Mathematics Instruction

	Component 1	Component 2	Component 3	Total
<u>Sample 1(n = 5,120)</u>				
Sum of squared loadings	4.74	4.00	3.07	11.80
Percentage of variance explained	14.8%	12.5%	9.6%	36.9%
Internal consistency reliability	0.84	0.79	0.72	—
Number of items	13	11	8	32
<u>Sample 2(n = 5,050)</u>				
Sum of squared loadings	4.84	3.88	2.97	11.69
Percentage of variance explained	15.1%	12.1%	9.3%	36.5%
Internal consistency reliability	0.84	0.78	0.72	—
Number of items	13	11	8	32

Note. Internal consistency reliability is for an equally weighted composite of all variables with loadings greater than or equal to .40 in the analysis of Sample 1.

Table 4.7. Final Principal Component Analysis Results for Mathematics Instruction

Instructional practices	Components		
	Number sense instruction	Student-centered mathematics instruction	Computation instruction
Ordering objects	.68	.27	.00
Making/copying patterns	.64	.14	.14
Name geometric shapes	.64	.11	.13
Sort into subgroups using rule	.64	.31	-.02
Counting by 2s, 5s, and 10s	.60	.02	.19
Write numbers 1 to 10	.60	.00	-.07
Relation between number and quantity	.58	.06	.03
Write all numbers 1 to 100	.52	.11	.17
Identify relative quantity	.52	.20	.23
Recognizing ordinal numbers	.51	.15	.29
Reading simple graphs	.48	.23	.27
Count out loud	.47	.11	.10
Simple data collection/graphing	.46	.29	.22
Solve math with partner	.09	.72	.08
Work on problems with several students	.11	.68	.11
Mixed group math work	.06	.63	.03
Solve real-life math problems	.13	.63	.24
Peer tutoring	.07	.57	.13
Explain/solve math problems	.08	.51	.35
Math-related games	.28	.51	-.03
Movement to learn math	.25	.49	-.04
Music to learn math	.23	.43	-.08
Use calculator	.09	.41	-.13
Adding two-digit numbers	.08	.12	.67
Subtracting two-digit numbers	.04	.13	.63
Subtracting single-digit numbers	.30	.04	.57
Routine practice or drill	.01	-.13	.54
Do math worksheets	.00	-.17	.52
Reading two-digit numbers	.29	.08	.49
Place values	.25	.12	.48
Use math for word problems	.17	.29	.46
Sum of squared loadings	4.74	3.81	3.04
Percentage of variance explained	15.28	12.28	9.80
Internal consistency reliability	.84	.79	.72
Number of items	13	10	8

Note. Internal consistency reliability is for an equally weighted composite of all variables with loadings greater than or equal to .40 on each particular component.

Summary

The two principal component analyses yielded a total of six instructional practices, three for reading and three for mathematics. The three-factor solutions for both subject areas explain an acceptable level of variance—36.7% and 37.4% for reading and mathematics, respectively. Additionally, each component in the final factor solutions for both reading and mathematics possesses a relatively high degree of internal consistency, indicating the overall strength and robustness of the measures. Interestingly, the resultant factors for reading and mathematics are conceptually parallel and somewhat hierarchical in nature. For example, letter sense and number sense instructional practices comprise the necessary basic skills to gain proficiency in fluency and comprehension in reading and computation for mathematics, respectively. Arguably, the student-centered instructional practice variables for both subjects involve instructional activities that engage higher-order skills and promote construction of meaning—a key goal of many reform-oriented practices. The next issue to be considered is the extent to which these instructional practice variables are related to the student achievement measures, and how usage of these instructional practices varies by a teacher’s certification status. Before addressing that issue it is important to note that these measures are indicative of teachers’ espoused practices rather than their actual practices. The measures are reliable only to the extent that teachers accurately reported them. Readers are advised to interpret the findings with this caveat in mind. Limitations of teachers’ self-report data about their instructional practices is discussed in Chapter 5.

Correlation Analysis of Teacher Instructional Practice Variables and Student Achievement Scores

The correlation analyses explore the extent to which students' mathematics and reading IRT scores are related to various instructional practices.

Reading Instructional Practices

Table 4.8 presents the association of spring first grade reading scores and letter sense, fluency and comprehension, and student-centered literacy instruction. Table 4.8 shows that letter sense instruction ($r = -.23$) and student-centered literacy instruction ($r = .14$) have statistically significant relationships with achievement ($p < .01$) but that these relationships were generally weak. The use of student-centered literacy instruction is associated with increases in reading scores, while the use of letter sense instruction is associated with decreases in reading scores.

Mathematics Instructional Practices

The correlation analysis performed for the mathematics instructional practice variables shows relatively weaker associations with the spring first grade mathematics score. Table 4.9 presents the association of spring first grade mathematics scores and number sense, computation, and student-centered mathematics instruction. The results indicate weak relationships, though some are statistically significant, between instructional practices and achievement. Number sense instruction ($r = -.07$) is negatively associated with the spring mathematics scores. Computation ($r = .08$) and student-

centered mathematics instruction ($r = .07$) are positively associated with spring first grade mathematics scores.

Table 4.8. Summary of Correlation Coefficients for Reading Instructional Practice Variables and First Grade Reading Achievement (N = 3,151)^a

	1	2	3	4
(1) First Grade Reading Achievement, R_IRT_4Z	—	-.23**	.01	.14**
(2) Letter Sense Instruction, ZREAD_FR		—	-.04*	-.02*
(3) Fluency and Comprehension Instruction, ZREAD_FC			—	.05**
(4) Student-Centered Literacy Instruction, ZREAD_CR				—

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user’s file).

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4.9. Summary of Correlation Coefficients for Mathematics Instructional Practice Variables and First Grade Mathematics Achievement (N = 3,151)^a

	1	2	3	4
(1) First Grade Mathematics Achievement, M_IRT_4Z	—	-.07**	.08*	.07**
(2) Number Sense Instruction, ZMATH_MR		—	.001	.04*
(3) Computation Instruction, ZMATH_C			—	.00
(4) Student-Centered Mathematics Instruction, ZMATH_SC				—

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user’s file).

* $p < .05$, ** $p < .01$, *** $p < .001$

In sum, the instructional practice variables in both reading and mathematics yielded statistically significant but weak associations with achievement test scores. The direction of association will be useful in understanding whether a particular strategy is related to increasing or decreasing student achievement. These relationships, of course, are only bivariate; the regression analysis provides further clarification concerning the effects of each variable on student achievement, controlling for related factors.

Descriptive Analyses of Teachers' Use of the Identified Instructional Practices by Certification Status

Tables 4.10 and 4.11 provide descriptive statistics for teachers' usage of the identified instructional practices by certification status. Each of the instructional practice variables is expressed in a z-score metric, in which mean differences can be interpreted as a percentage of an SD in the distribution of the given instructional practice across all students. One-way ANOVAs with Tukey HSD post hoc contrasts (emergency vs. standard and advanced vs. standard) were used to determine whether there is a relationship between certification status and student characteristics. The results of these tests are indicated in the tables.

Reading Instructional Practices

One-way ANOVAs were conducted on each of the three reading practice factors to determine whether there were mean differences between teachers in the three certification groups. Means and standard deviations for the three groups are presented in Table 4.10. Of the three reading instructional practice variables, only the mean difference for student-centered literacy instruction is statistically significant ($p < .05$) among the

three certification statuses. Emergency certified teachers utilize student-centered literacy instruction ($p < .05$) most frequently (.12 SD vs. -.02 SD), followed by teachers with advanced certification (.07 SD vs. -.02 SD). Tukey HSD tests indicated that teachers with standard certification were significantly ($p < .05$) less likely to use student-centered instruction than teachers with either emergency or advanced certification.

Table 4.10. Descriptive Statistics for Reading Instructional Practice Variables by Teacher Certification Level (N = 3,151)^a

	Emergency certified teachers (8.69%)	Standard certified teachers (82.2%)	Advanced certified teachers (9.1%)
Letter Sense Instruction, ZREAD_FR	.05	.00	-.02
Fluency and Comprehension Instruction, ZREAD_FC	.02	.00	.01
Student-Centered Literacy Instruction, ZREAD_CR ^{b*}	.12*	-.02	.07*

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K-1) (public-user's file).

^b Significance based on one-way ANOVA with Tukey HSD follow-up tests (Emergency vs. Standard; Standard vs. Advanced).

* $p < .05$, ** $p < .01$, *** $p < .001$

Mathematics Instructional Practices

The tests of independent significance for the mathematics instructional practices described in Table 4.11 indicate that only the mean difference for number sense instruction is statistically significant ($p < .01$). Emergency certification status teachers are most likely to utilize number sense instruction (.20 SD vs. -.01 SD), while teachers with

advanced certification are least likely to use number sense instruction (-.08 SD vs. -.01 SD). Tukey HSD follow-up tests indicate that teachers with standard certification are significantly ($p < .01$) less likely to use number sense instruction than either emergency or advanced certification status teachers.

Table 4.11. Descriptive Statistics for Mathematics Instructional Practice Variables by Teacher Certification Level (N = 3,151)^a

	Emergency certified teachers (8.69%)	Standard certified teachers (82.2%)	Advanced certified teachers (9.1%)
Number Sense Instruction, ZREAD_FR ^{b*}	.20**	-.01	-.08**
Computation Instruction, ZREAD_FC	-.03	.01	-.03
Student-Centered Mathematics Instruction, ZREAD_CR	-.04	.00	.05

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K–1) (public-user’s file).

^b Significance based on one-way ANOVA with Tukey HSD follow-up tests (Emergency vs. Standard; Standard vs. Advanced).

* $p < .05$, ** $p < .01$, *** $p < .001$

Summary

In sum, teachers’ use of the identified instructional practices varies little across certification status. The mean differences are significant for only two instructional practices—student-centered literacy instruction and number sense instruction. Emergency certified teachers utilize these strategies most often. Teachers with advanced certification are least likely to use number sense instruction, and only a little less likely to use student-

centered literacy instruction than teachers with standard certification. The absence of statistical significance for the mean differences of the other instructional practices suggests either measurement error or that certification status has relatively little effect on a teacher's decision to utilize a particular strategy. Recalling the correlations from the previous section above, emergency certification status teachers, interestingly, tended to use the strategy most associated with reform-oriented instruction in reading (student-centered literacy instruction), though in mathematics they were no more likely to use the comparable reform-oriented set of practices (student-centered mathematics instruction) than were other teachers.

Results of OLS Regression Analyses

Tables 4.12 and 4.13 summarize the OLS regression analyses for first grade reading and mathematics achievement scores, respectively. While prior tables presented only bivariate statistics, these tables present multivariate statistics. Each table includes the results for five progressively more complex models. Model 1 is the simplest model and presents an unadjusted baseline estimate of the association between certification status and end-of-first-grade achievement. The coefficient for emergency certification status is the estimate of the mean difference in achievement between students taught by a teacher with emergency certification and those taught by a teacher with standard certification; similarly, the coefficient for advanced certification status is the estimate of the mean difference in achievement between students taught by a teacher with advanced certification and a teacher with standard certification. Because the dependent variable, end-of-first-grade achievement, is standardized, the coefficient can be interpreted as an

“effect size”—that is, the percentage of an SD increase (or decrease) in achievement associated with a unit increase in the independent variable.

Subsequent models enter additional variables associated with teacher qualifications, student characteristics, teacher practices, and possible interaction effects between certification status and other variables in the model. Model 2 enters three teacher characteristics—teacher race (nonwhite teacher), degrees (master’s and beyond vs. less than master’s) and years of experience (less than 2 years vs. 2–6 years and more than 6 years vs. 2–6 years). Model 3 enters student demographic variables (nonwhite vs. white and socioeconomic status), students’ end-of-kindergarten achievement scores, and a control for test administration differences between students in the time that elapsed between the end-of-kindergarten and the end-of-first-grade testing. Model 4 enters the teacher practice factors described in Tables 4.5 and 4.7.

Model 5, the final and most complex of the models, tests for possible interaction effects between teachers’ certification status and other variables in the model. To examine the possibility of interaction effects, interaction terms were calculated with certification status for all student characteristic variables entered in Model 3 and for all teacher practice variables entered in Model 4. The interaction terms (14 in all) were then entered into the model using a stepwise procedure, entering terms with the largest coefficient first and restricting entry to only those terms with a statistically significant coefficient ($p < .05$). Results of the stepwise procedure are reported in Model 5. The corresponding R^2 for

Table 4.12. Summary of OLS Regression Analysis for Variables Predicting Average First Grade Reading IRT Scores (N = 3,151)^a

	Model 1		Model 2		Model 3		Model 4		Model 5	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Certification Status										
Emergency Teacher	-.284***	.066	-.196**	.070	-.034	.046	-.030	.045	-.012	.046
Advanced Teacher	-.065	.065	-.059	.066	.037	.043	.028	.043	.027	.043
Teacher Characteristics										
Nonwhite Teacher			-.273**	.052	-.085*	.035	-.079*	.035	-.077*	.035
Master's and Beyond			-.018	.039	-.029	.025	-.014	.025	-.016	.025
<2 Years of Teaching			-.101	.069	-.023	.045	-.038	.044	-.042	.044
>6 Years of Teaching			.034	.050	.040	.033	.021	.032	.022	.032
Student Characteristics										
Nonwhite Student					-.146***	.029	-.145***	.029	-.144***	.029
SES					.050***	.013	.056***	.013	.055***	.013
K-Reading Score					.694***	.013	.679***	.013	.670***	.014
Elapsed Time					.091***	.012	.091***	.012	.091***	.012
Instructional Practices										
Letter Sense Instruction							-.062***	.012	-.062***	.012
Fluency and Comprehension							.068***	.012	.069***	.012
Student-Centered Literacy Instruction							-.027*	.012	-.025*	.012
Interactions										
Emergency Certification by K-Reading Score									.117**	.044
<i>R</i> ²	.007		.018		.582		.592		.593	

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K-1) (public-user's file).

Note. The following measures are expressed as z-scores with a mean of 0 and an SD of 1: SES, K-Reading Score, Letter Sense Instruction, Fluency and Comprehension Instruction, Student-Centered Literacy Instruction.

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4.13. Summary of OLS Regression Analysis for Variables Predicting Average First Grade Mathematics IRT Scores ($N = 3,151$)^a

	Model 1		Model 2		Model 3		Model 4		Model 5	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Certification Status										
Emergency Teacher	-.229**	.067	-.196**	.071	.005	.045	.010	.044	.008	.044
Advanced Teacher	-.106	.066	-.109	.067	-.018	.042	-.020	.042	-.018	.042
Teacher Characteristics										
Nonwhite Teacher			-.303***	.053	-.054	.034	-.050	.034	-.050	.034
Master's and Beyond			.035	.039	.015	.025	.021	.024	.019	.024
<2 Years of Teaching			.038	.070	.048	.044	.061	.043	.063	.043
>6 Years of Teaching			-.007	.051	-.019	.032	-.023	.031	-.020	.031
Student Characteristics										
Nonwhite Student					-.127***	.029	-.135	.029	-.135**	.029
SES					.076***	.013	.079***	.013	.077***	.013
K-Reading Score					.719***	.013	.706***	.013	.698***	.014
Elapsed Time					.082***	.011	.081***	.011	.081***	.011
Instructional Practices										
Number Sense Instruction							-.047***	.011	-.047***	.011
Computation Instruction							.086***	.011	.086***	.011
Student-Centered Math Instruction							.017	.011	.018	.012
Interactions										
Adv. Certification by K-Math Score									.112*	.042
Emrg. Certification by K-Math Score									.116*	.045
Emrg. Certification by Number Sense									-.099*	.041
R^2	.005		.017		.616		.627		.628	

^a Source: National Center for Education Statistics, Early Childhood Longitudinal Study (K-1) (public-user's file).

Note. The following measures are expressed as z-scores with a mean of 0 and an SD of 1: SES, K-Mathematics Score, Number Sense Instruction, Computation Instruction, Student-Centered Mathematics Instruction.

* $p < .05$, ** $p < .01$, *** $p < .001$

each model is reported at the bottom of the table; it represents the proportion of the total variance in end-of-first-grade achievement “explained by” or “accounted for” by a model.

Five variables included in earlier tables were excluded from the final models presented in Tables 4.12 and 4.13 (students’ gender, teachers’ age, average class size, school’s Title I status, and school’s minority enrollment). Excluded variables had no statistically significant relationship to achievement in the final models for mathematics or reading. Teachers’ degrees and experience were retained in the models even though they failed to achieve statistical significance, because these variables are typically included in studies of teacher certification status.

Reading Scores

Model 1 indicates that students taught by teachers with emergency certification had lower end-of-first-grade reading achievement scores than students taught by teachers with standard certification, the difference being equivalent to slightly more than one-quarter of a SD in end-of-first-grade reading achievement ($b = -.284$ SD). Although students taught by teachers with advanced certification also scored, on average, lower than students of teachers with standard certification, the difference is not statistically significant. When teacher characteristics are entered in Model 2, they “explain” only a small part of the difference in achievement scores between students of teachers with emergency and standard certification. The coefficient is slightly strengthened ($b = -.284$ vs. $-.196$ SD), and only the coefficient for nonwhite teacher is statistically significant.

The main explanation for the difference in achievement scores is associated with differences in the characteristics of students taught by teachers with emergency certification and standard certification (see Model 3). Once these characteristics are taken into consideration, the mean difference in achievement is nonsignificant. End-of-first-grade achievement scores are lower for nonwhite students than for white students ($b = -.146$ SD), higher for higher SES students than lower SES students ($b = .050$ SD), and higher for students who had more time to learn between testing than students who had less time to learn ($b = .090$ SD). The strongest effect, however, is students' prior knowledge. Students who had stronger reading skills entering the first grade also had stronger reading skills at the end of first grade ($b = .694$ SD). The magnitude of the effect associated with prior knowledge increases the R^2 to .582 from less than 1% in the previous models.

Model 4 enters the teacher practice variables. These variables do not noticeably alter the coefficients associated with other variables in the model, but each is related to end-of-first-grade reading achievement. The greater the emphasis on fluency and comprehension, the greater students' achievement level ($b = .068$ SD), whereas the greater the emphasis on letter sense and student-centered literacy instruction, the lower students' achievement level ($b = -0.62$ and $-.027$ SD, respectively). Although these effects are relatively small by conventional standards, the effects for letter sense and the effects for fluency and comprehension are as large, if not larger than, the effects for SES (controlling for prior knowledge).

Model 5 tests for possible interactions. Only one interaction term proved statistically significant: the interaction between the variables for teachers' certification

status and students' end-of-kindergarten achievement. Although emergency certification status remains nonsignificant in this model, it has, nonetheless, an effect on the relationship between students' prior knowledge and end-of-first-grade achievement. Prior knowledge is a more powerful predictor of achievement for students taught by teachers with emergency certification ($b = .670 + .117 = .787$ SD) than for students taught by teachers with standard certification ($b = .670$ SD). One possible interpretation of the interaction effect is that emergency certified teachers are less able to narrow the achievement gap between students who enter first grade with varying reading skills.

Mathematics Scores

Model 1 indicates that students taught by teachers with emergency certification had lower end-of-first-grade mathematics achievement scores than students taught by teachers with standard certification, the difference being equivalent to less than a quarter of an SD in end-of-first-grade mathematics scores ($b = -.229$ SD). Similarly, students of teachers with advanced certification scored lower on average than students of teachers with standard certification, but the difference is not statistically significant. When teacher characteristics are considered in Model 2, they “explain” slightly more of the difference in achievement scores between students of teachers with emergency and standard certification ($b = -.229$ vs. $-.196$ SD), though only nonwhite teacher is statistically significant.

Like the OLS models for reading scores, the primary explanation for the differences in achievement scores is associated with the student characteristics variables entered in Model 3. The mean difference in mathematics achievement is nonsignificant

once these characteristics are taken into account. End-of-first-grade achievement scores are lower for nonwhite students than white students ($b = -.127$ SD), higher for higher SES students than lower SES students ($b = .076$ SD), and higher for students who had more time to learn between testing than students who had less time ($b = .082$ SD). As with reading scores, the strongest effect is students' prior knowledge. Students entering the first grade with strong mathematics skills also had stronger mathematics skills at the end of the first grade ($b = .719$ SD).

Model 4 enters the instructional practice variables. These variables alter the coefficients associated with other variables in the model only slightly, but two of the three instructional practice variables—number sense instruction and computation instruction—are related to end-of-first-grade mathematics achievement. The greater the emphasis on computation instruction ($b = .086$ SD), the higher the students' achievement level, whereas the greater the emphasis on number sense instruction, the lower the students' achievement level ($b = -.047$ SD). Student-centered mathematics instruction has a small and statistically insignificant, but positive, effect on mathematics achievement ($b = .017$ SD). Again, these effects are relatively small, though the effect for computation instruction is comparable to the effect of SES (controlling for prior knowledge).

Model 5 tests for interaction effects. Each of the three interaction terms proved statistically significant. Two of the interactions involve the effect of certification status (i.e. emergency certification and advanced certification) and students' prior knowledge in mathematics. The third interaction tests the effects of certification status and number sense instruction. Although certification status remains nonsignificant in this model, it has, nonetheless, an effect on the relationship between students' prior knowledge and

end-of-first-grade achievement. Prior knowledge is a more powerful predictor of student achievement for students taught by teachers with emergency certification ($b = .698 + .116 = .814$ SD) and advanced certification ($b = .698 + .112 = .810$ SD) than for students of teachers with standard certification ($b = .698$ SD). As with reading instruction, one possible interpretation of the interaction effect is that emergency and, surprisingly, advanced certified teachers are less able to narrow the achievement gap between students who enter the first grade with varying mathematics skills; the direction of the interactions suggests that these teachers may engage in practices that benefit students who enter the first grade with higher levels than lower levels of mathematics knowledge. The final interaction term tests the effect of emergency certification and number sense instruction ($b = -.099 -.047 = -.146$ SD). Although number sense instruction is nonsignificant in this model, it does, nonetheless have an effect on student achievement when an emergency certified teacher utilizes it. This effect, in fact, appears more deleterious than the effects of either race or SES on mathematics achievement. One interpretation of this interaction is that the choice of instructional practices matters more for mathematics achievement in the case of emergency certified teachers.

Summary

The results from the two OLS regression analyses are quite similar. In both reading and mathematics, emergency certification status is significant only in the non-adjusted models. Emergency certification status remains statistically significant when teacher characteristics are controlled in the model for reading, but not for mathematics. In subsequent models, which control for student characteristics and instructional practices,

emergency and advanced certification status is nonsignificant and not associated with either reading or mathematics achievement. Teacher characteristics—time on the job and having a master’s-level education or beyond—did not affect student achievement in either subject area; however, having a nonwhite teacher resulted in slightly lower reading achievement. Prior knowledge is a powerful predictor in both reading and mathematics achievement. Prior knowledge “explains” nearly 67% of the variance in reading achievement and 69% of the variance in mathematics achievement. With respect to the instructional practice variables, fluency and comprehension instruction and computation instruction were statistically significant in all models and associated with gains in reading and mathematics achievement, respectively; but letter sense instruction and student centered instruction in reading and computation instruction in mathematics were statistically significant and negatively associated with achievement.

The interaction terms for both reading and mathematics underscore the complexity of and highlight important nuances regarding the effects of certification status on student achievement. In reading and mathematics, for example, the effect of the relationship between students’ prior knowledge and their teachers’ emergency certified status reveals that teachers with emergency certification may actually be less effective in effecting gains in student achievement, particularly for lower achieving students. A similar conclusion may apply to teachers with advanced certification status in mathematics. Moreover, the selection of instructional practices by emergency certified teachers matters in students’ subsequent mathematics achievement. These findings and those from the previous analyses are discussed in greater depth in chapter 5.

CHAPTER V: DISCUSSION AND CONCLUSION

This chapter is organized around the eight major findings of this study. The chapter begins with a review and discussion these findings. Following this review, the three research questions that guided the study are individually addressed. Then, arguments are presented for why certification does or does not matter based on these data, followed by a discussion of the implications for the continued use of certification status as an indicator of teacher quality. The chapter concludes with an assessment of the limitations of the study and recommendations for future research.

Review and Discussion of Major Findings

Finding 1: The characteristics of first grade public school students vary by their teacher's certification status.

The analysis of descriptive and inferential statistics of student characteristics by teacher certification type indicated that there were clear and statistically significant differences between the types of students taught by teachers in the sample with emergency, standard, and advanced certification. Although no significant differences were found in student gender or the time lapse between kindergarten and first grade test administration, significant differences existed in every other tested characteristic—race, SES, and achievement test scores. Consistent with findings by others (Darling-Hammond, 2000; Laczko-Kerr & Berliner, 2002; Qu & Becker, 2003), emergency certified teachers in this sample were assigned the most at-risk students. Specifically, the

students of emergency certified teachers tended to be poorer and nonwhite, and scored below the mean on both kindergarten and first grade reading and mathematics assessments. Comparatively, students assigned to teachers with advanced certification had fewer at-risk characteristics. These students were mostly white, had higher SES, and scored above the mean on reading and mathematics assessments both in kindergarten and first grade. One interesting finding was that students assigned to teachers with advanced certification had (statistically significant) slightly lower kindergarten and first grade reading and mathematics scores than students of teachers with standard certification. This observation requires more sophisticated testing than this analysis provided; however, as suggested by the subsequent analyses in Chapter 4, there may be a point of diminishing returns in terms of increased student achievement outcomes associated with certification past the standard level.

Finding 2: The demographic characteristics of first grade public school teachers vary by certification status.

The inferential statistics used as part of the preliminary analysis of teacher characteristics indicated statistically significant differences in all the teacher characteristics studied in the analysis—race, age, degrees, and teaching experience. Emergency certified teachers tended to be younger, nonwhite, less experienced, and least likely to possess a master’s degree (in any field). Conversely, teachers with standard and advanced certification were more likely to be older, white, more experienced, and possess a master’s degree.

Murnane and Phillips (1981) found that teachers with 15 years of experience or more had a positive effect on student achievement at the elementary level. Ferguson and

Ladd (1996), however, found no significant differences in student achievement after teachers had had at least 5 years of teaching experience. The proportion of teachers in the current sample that have more than 6 years of experience to those with 2 or fewer years of experience is nearly 3:1. These analyses, however, failed to identify an independent effect of teachers' experience, perhaps, in part, because most of the teachers in the sample had five or more years of experience.

Teachers' advanced degrees do not seem to have a statistically significant effect on student achievement. Rice's (2003) review of the literature related to teacher degrees examined six empirical studies on the subject, which on balance showed small effects of advanced degrees on black students. The OLS regression models for reading and mathematics, discussed in Findings 7 and 8 also demonstrate that teachers' possession of an advanced degree is not associated with students' reading or mathematics achievement levels.

Finding 3: The characteristics of first grade public school teachers' schools vary by the teachers' certification status.

The ANOVA for class size indicated that there was no significant difference in class size among the three groups of teachers. Other tests indicated that as minority enrollment decreases, the percentage of teachers with standard and advanced certification increases. A similar pattern was observed with respect to receipt of Title I funds, an indicator of student poverty in schools. These findings indicate that the schools that teachers with emergency certification teach in are most likely to serve students who are at risk, and the converse is true for teachers with advanced certification.

Finding 4: The principal component analysis generated reliable instructional practice variables.

The principal components analysis (PCA) generated a set of six of instructional practice variables—three for reading instruction and three for mathematics instruction. These variables were used in a correlation analysis to determine which instructional practices are related to mathematics and reading achievement. Additionally, these factors were used in the regression analyses to determine the effect of each on mathematics and reading achievement respectively. To review, the mathematics practice variables included (a) student-centered mathematics instruction, (b) computation instruction, and (c) number sense instruction. Each of the mathematics factors had a reliability coefficient greater than or equal to .72, and they explained a combined 37.37% of the variance for all mathematics activity items. The reading instructional practice variables included (a) student-centered reading instruction, (b) fluency and comprehension instruction, and (c) letter sense instruction. Each of the reading factors had a reliability coefficient greater than .77, and they explained a combined 36.45% of the variance for all reading activity items. Given the stability of these measures, they were used to determine the extent to which use of these practices varied according to teachers' certification status, discussed in Finding 5, as well as to determine the extent to which these use of these practices are related to student achievement in the OLS regression models, discussed in Findings 6 and 7.

Finding 5: Several of the instructional practice variables demonstrated weak but statistically significant associations with student achievement in first grade mathematics and reading.

The correlational analysis performed for the reading instructional practice variables shows a generally weak correlation with student achievement. The use of student-centered literacy instruction was associated with increases in reading scores, while the use of letter sense instruction was associated with decreases in reading scores. Fluency and comprehension instruction scores were not related to student reading achievement ($r = .016$).

The correlation analyses found weak but statistically significant relationships between each of the mathematics instructional practice variables and end-of-first-grade mathematics scores. Specifically, student-centered mathematics and computation instruction were associated with higher end-of-first-grade mathematics scores. Number sense instruction is negatively associated with mathematics achievement. These results indicate that teachers increase their likelihood of effecting gains in student achievement in mathematics by focusing on the former practices.

In summary, this analysis revealed statistically significant, but generally weak, associations between the instructional practice variables and student achievement in both reading and mathematics.

Finding 6: Teachers of varying certification statuses differed little in their use of the examined mathematics and reading instructional practices.

ANOVAs for each instructional practice variable were conducted to determine whether differences existed among the three types of teachers in terms of instructional practice. With respect to the reading practices, there was a significant difference between teachers with emergency and standard certification in their use of student-centered literacy instruction. Teachers with standard certification utilized this instructional practice less frequently than did teachers with emergency certification, and teachers with advanced certification also used this practice more frequently than did teachers with standard certification. No other differences, however, were statistically significant, indicating that there is virtually no variation in the frequency with which teachers with emergency, standard, or advanced certification utilize these strategies. The ANOVAs for the mathematics instructional practices indicate that the only significant difference among the three groups is in the use of number sense instruction by teachers with emergency and standard certification. As with the finding related to reading, this finding indicates that there is little variation in the frequency with which teachers with emergency, standard, or advanced certification utilize these strategies.

Overall, certification status does little to differentiate the frequency with which these teachers use the examined instructional practices. The notable exceptions are the differences between teachers with emergency and standard certification with respect to student-centered literacy instruction and number sense instruction.

This finding requires additional explanation. First, it is important to acknowledge that a teacher's choice to engage in any of these instructional activities is mediated by

several constraints—namely, the curriculum (Cohen & Hill, 2000), the teacher’s knowledge base and experience (Shulman, 1987), and the students’ needs and prior knowledge (Tomlinson, 2001; Marzano, 1992). Emergency certified teachers are most likely to teach students with lower levels of reading achievement, which leads to lower-level reading approaches, such as letter sense instruction, for these children (Coles, 2003). Given this context, one plausible explanation for the differences in the use of various instructional practices by teachers with emergency and standard certification is that these teachers are simply using the strategies dictated by adopted curricula in response to actual and perceived student need. Second, by its nature, student-centered instructional approaches are more difficult to implement and require the teacher to draw from a more sophisticated instructional repertoire than do the other instructional strategies examined. Some research (Brophy & Good, 1986; Clandinin & Connelly, 1987, 1995, 1996; Russell & Munby, 1991) indicates that a teacher’s instructional repertoire results from a combination of pedagogical development, time on the job, and subject-matter knowledge. These data do not allow for a clear understanding of the dynamic context in which teachers make instructional decisions. Therefore, this finding should be interpreted with caution. However, as discussed in Finding 7, another important aspect of teacher practice may be the possibility of differential effects when used by teachers with emergency certification status.

Finding 7: The OLS regression analysis yielded several statistically significant predictors of first grade reading achievement.

Analysis using OLS regression indicated that the chief predictors of first grade reading achievement were being a white student with above average SES and above

average kindergarten reading scores, and having teachers who used fluency and comprehension strategies. By far the strongest predictor of first grade reading achievement was kindergarten reading scores. The effects of prior learning or prior success on such assessments as predictors of future success have been well documented (see Rowan, Correnti, & Miller, 2002, for example).

Certification status seemed to predict reading achievement in an uncontrolled model, but after controlling for teacher characteristics—nonwhite teacher, teacher experience (2 or fewer years of experience and 6 or more years of experience), and having a master's-level education or beyond, the effect disappeared. A small, statistically significant, negative effect on end-of-first grade reading scores for students of nonwhite teachers is apparent. The existing literature on the significance of teacher's race and student achievement is inconclusive and may be an area that warrants further research to determine whether teachers' race truly matters.

The interaction term computed for emergency certification and spring kindergarten reading score (.117 SD, $p < .01$) suggested that the effect of having an emergency certified teacher serves to widen the achievement gap. The implications of this gap are discussed in the next section of this chapter.

The reading instructional practice variables, which have stronger effects than certification status and key student characteristics such as SES⁶, predicted reading achievement in positive and negative ways. Both letter sense instruction and student-

⁶It is important to note that prior achievement (i.e. kindergarten achievement) captures some of the prior effects of SES on first grade achievement. Nonetheless, teachers can only influence learning during the time that they have students in their classes – that is, the effects of these practices and SES are roughly the effects during the time that first grade teachers have to influence achievement. The statistical method used here cannot mediate these endogenous qualities of SES. Therefore, readers should interpret the reported SES partial regression coefficients with this caveat in mind.

centered literacy instruction negatively predicted increased reading scores. Fluency and comprehension instruction, in contrast, positively predicted increased student achievement. Given the emphasis on letter sense instruction and the appeal of student-centered approaches, these findings seem to argue against many favored reforms (basic and constructivist). A few explanations are plausible for this finding. First, this type of statistical analysis is not well suited to understanding the nuances of how teachers marshal various resources to effect gains in student achievement. Rather, the results are at best exploratory, and at worst a crude indicator of likely strategies teachers could use to enhance students' learning in reading. Second, the results presented here may actually reflect issues with the measure of student achievement used in this analysis. The ECLS-K child assessment is a standardized test that seeks to measure a student's ability to read, study pictures, and make general interpretations of literature. In this sense, what might be measured here is instruction targeted at doing well on some other assessment or set of standards, rather than knowledge and skills assessed by the ECLS-K assessment.

Finding 8: The OLS regression analysis yielded several statistically significant predictors of first grade mathematics achievement.

The regression analysis for mathematics achievement indicated that the effects of teachers' certification status on students' mathematics achievement were nonsignificant after controlling for teacher experience (2 or fewer years of experience and 6 or more years of experience) and having a master's-level education or beyond. The best predictors of first grade mathematics achievement were being a white student with an above mean SES and an above average kindergarten mathematics test score, and having a

teacher who uses computation instruction. The strongest predictor of first grade mathematics achievement was kindergarten mathematics score. Other researchers (Dossett & Munoz, 2000; Munoz, 2000; Fidler, 2000) have found similar patterns with early-grade mathematics achievement. Munoz (2000) did find, however, that a teacher's certification status—possession of a standard certificate—had a small effect on students' mathematics achievement.

Computation instruction and number sense had statistically significant effects in Model 4, which controlled for certification status and student and teacher characteristics; computation instruction had a positive effect, while number sense had a negative effect. Student-centered mathematics instruction had no statistically significant effect in any model. Given the emphasis on developing arithmetic skills at this grade level, the association with computation instruction makes sense intuitively. Considering the instructional reforms touted by standards-based reformers in mathematics (the NCMT, for example), the failure of student-centered instruction to register an effect is noteworthy. These findings seem similar to those of Klein et al. (2000), discussed earlier, who found that teachers generally supported student-centered approaches in general but did not apply them in the classroom. In other words, the absence of any effect may be due to teachers saying that they use these practices when in actuality they do not.

In this model, three indirect effects on end-of-first-grade mathematics achievement are noteworthy. The first indirect effect involved emergency certified teachers and kindergarten mathematics scores. The second indirect effect involved teachers with advanced certification. In both cases, the effect of certification status (emergency and advanced) enhances the effect of prior mathematics achievement. The

interpretation of this effect with regard to emergency certified teachers is that they tend to increase the achievement gap between students who enter first grade with higher and lower achievement levels. The interpretation with regard to teachers with advanced certification is similar: These teachers also are unable to close the achievement gap, seemingly benefiting higher achieving students over lower achieving students. The third indirect effect involved the interaction computed with emergency certification and number sense, and indicates that emergency certified teachers who use this strategy are least likely to effect gains in end-of-first-grade mathematics scores using this strategy. This finding is related to Finding 6, the correlational analysis, which indicated that teachers with emergency certification were also the most likely of the three groups of teachers to use this strategy. The significance of these indirect effects is discussed in greater depth later in this chapter in the Does Certification Matter? section.

Examination of Research Questions

Research Question 1

The first research question of this study was: How do the characteristics of first grade public school teachers, their schools, and their students vary by certification status? Results indicated that first grade public school teachers with different certification statuses differed on every dimension under study, with the exception of student gender and class size. Evidence from the descriptive analysis indicated that emergency certified teachers are most likely to be white and teach nonwhite students with lower SES and below average kindergarten and first grade IRT scores, in high minority enrollment schools that receive Title I funds. The sample of emergency certified teachers had a

higher percentage of minority teachers. Though the differences between teachers with standard and advanced certification are significant, their demographic profiles were virtually the same. These teachers tended to be white and had students with above average kindergarten and first grade IRT scores and above mean SES, and taught in schools with lower minority enrollments and less receipt of Title I funds. On the whole, students of teachers with standard certification had higher mathematics and reading IRT scores than students of teachers with advanced or emergency certification.

Research Question 2

The second research question was: Does a teacher's certification status predict the frequency of use of various classroom instructional activities in reading and mathematics? Results indicated that little differentiation existed in the usage of the identified instructional practices in reading and mathematics based on certification status. The evidence suggests either that teachers use these practices consistently, showing little variability in use, or that these practices are highly related. Where small but statistically significant differences existed, the evidence suggests that emergency certified teachers were least likely to use instructional practices associated with higher reading or mathematics scores.

Research Question 3

The third research question was: Does a teacher's certification status predict how students will perform in reading or mathematics, or both? The results indicated that a teacher's certification status is not a predictor of how students will perform in either

reading or mathematics. The analysis suggests that emergency certification has a statistically significant but weak negative effect on student achievement in reading and mathematics before other teacher attributes such as master's-level education or above and teacher experience (less than 2 years or greater than 6 years) are controlled for. When student-level controls are considered, however, the effects of certification status on reading and mathematics scores are nonexistent. The strongest predictors of first grade reading and mathematics achievement are prior reading and mathematics achievement, respectively. Although no direct effects of certification status were observed (after student and teacher controls), there were meaningful indirect effects of certification status, particularly in mathematics, with respect to prior learning and the use of certain instructional practices.

Does Certification Matter?

The findings discussed above highlight the complexity involved in understanding the benefits of certifying teachers. This study was undertaken to try to determine the extent to which certification status influences student achievement. The evidence presented above lends itself neither to a rallying cry for certification nor to an outright condemnation. Certification advocates can certainly point to the findings concerning the more or less indirect deleterious effects that emergency certification has on student achievement. However, those who question whether certification really matters can point to the findings that indicate that increasing a teacher's level of certification past the standard level seems to matter very little in terms of student outcomes (and might actually be detrimental for low-achieving students). The findings presented above show

that the case for or against certification is complex and underscore the nuances and difficulties involved in the issue of teacher certification.

This section presents two arguments, one for why certification does not matter and the other for why it does, based on the findings of this study. These arguments consider the direct and indirect effects of certification status on student achievement, salient instructional effects, and the overall value added by certification to student achievement. Consider the certification opponents' argument first, and then that of the proponents.

Why Certification Doesn't Matter

For opponents of certification, the evidence presented above can be used to make a persuasive case against certification. The findings suggest three reasons for opposing certification: (a) the direct effects of certification status on both reading and mathematics achievement are negligible; (b) instructional practices in reading and mathematics matter more than certification status for student achievement; and (c) advanced levels of certification add little to student achievement, especially in mathematics, where they may actually decrease achievement compared to standard certification. Consider each reason in greater depth below.

The direct effects of certification status on either reading or mathematics achievement are negligible. On its face, certification status does not seem to matter at all in terms of reading and mathematics achievement. In reading, the direct effects of certification status (e.g., the effect of standard certification vs. emergency certification) are absent once teacher and student characteristics are considered. For mathematics

achievement, the direct effects of certification status disappear when only teacher characteristics are considered. The absence of a certification effect in these data could result from measurement error or sampling error, but there is no indication that either occurred with these data. Other salient factors are more related to student achievement than certification status. One of these factors—instructional practices—is discussed below.

Instructional practices in reading and mathematics matter more than certification status for student achievement. The findings discussed above highlight the importance of the instructional practices that teachers utilize to affect students' learning. Although others (see Goldhaber & Brewer, 1996; Greenwald, Hedges, & Laine, 1996; Hanushek, 1986) have conducted similar analyses concerning the effects of teacher certification on student achievement, few have examined what teachers actually do in the classroom and how specific instructional practices might be related to teacher certification. These findings suggest that what teachers do in the classroom (i.e. the instructional practices they utilize) is critical to students' learning—more so than teachers' certification status. Some instructional practices are clearly better than others, at least in the first grade. Although the effects of prior achievement explain most of the variability in end-of-first-grade achievement, the effects of instructional practices are among the strongest in the models, comparable to or even slightly larger than the effects of SES after controlling for prior achievement.

Since first grade typically marks the beginning of a student's foray into academic subject matter, the findings from this study suggest that how subject matter is taught and who teaches it is critical to future performance, especially in mathematics. In the early

grades, students develop foundational skills necessary for mastering more complex subject matter in later grades. If certification status is an especially ineffective indicator of teacher quality, and only weakly related to what teachers do in the classroom, the focus should be on seeking or relying more heavily on alternatives, such as the quality of a teacher's actual instructional practices, especially in the early grades—a reportedly critical time in students' learning.

Advanced levels of certification add little to student achievement, especially in mathematics, where they may actually decrease achievement compared to regular certification. Given that teachers with advanced certification are likely to be the most experienced teachers and have the highest degrees, it is reasonable to expect students of these teachers to achieve more than the students of teachers with standard or emergency certification. This assumption, however, is not supported by these data. The descriptive statistics provided an initial indication that students of teachers with advanced certification scored higher than those of teachers with emergency certification, but lower (and below the mean) than students of teachers with standard certification, in both subject areas. Nor was there a direct effect of advanced certification status on reading or mathematics, even before entering teacher or student characteristics into the models.

Even more troubling, though, is the positive interaction between advanced certification status and students' prior mathematics achievement. Such an interaction indicates that students' prior knowledge is more important in determining what students learn about mathematics in classrooms taught by teachers with advanced certification than in classrooms taught by teachers with standard certification—in other words, teachers with advanced certification “add less value” to first grade mathematics

achievement than regularly certified teachers, particularly for students with lower levels of prior achievement. Why this might be true, and true only in mathematics, is difficult to determine. Teachers with advanced certification may place less emphasis on mathematics in the first grade, especially if their students enter first grade with above average knowledge, or they may use practices that are more effective with higher-knowledge students than lower-knowledge students. Neither explanation is comforting—especially the latter, which would imply a widening of the achievement gap between high- and low-achieving students.

In sum, the findings from this study support claims made by opponents of certification that certification policies are largely misguided. Policies that refuse to acknowledge alternative qualifications or that encourage or even mandate that teachers raise their certification status past the standard level may be a poor use of time and resources. This point is revisited, in greater depth, in the section on policy implications.

Why Certification Matters

To proponents of certification, the evidence from this study suggests that certification does indeed matter. Although the direct effects of certification status on both reading and mathematics achievement are negligible, the indirect effects of certification status present a compelling case for why certification matters. Proponents will emphasize the interaction between emergency certification and students' prior achievement, as well as the interaction between emergency certification status and specific instructional practices as reasons why certification matters. In each instance, the implication is that students taught by emergency certified teachers learn less in reading and mathematics

than students taught by teachers with standard certification. Consider these effects in greater depth below.

If certification status can differentiate between the worst and all other teachers, then there is some merit to certification. Teachers with emergency certification, like those with advanced certification, affect fewer gains in student achievement, particularly for lower-achieving students, but unlike teachers with advanced certification, they do so in both reading and mathematics. The interaction terms between emergency certification and prior achievement in reading and mathematics are virtually identical. One interpretation might be that emergency certified teachers focus less time on instruction than on other matters (e.g., classroom management or administrative requirements). A second interpretation might be that they are simply less effective at teaching, especially for students who enter first grade with less than average knowledge of reading and mathematics. This interpretation could indicate that emergency certified teachers, like advanced certified teachers, also add less value to student achievement and are not able to close the achievement gap in their classrooms.

Admittedly, the evidence presented in this study indicates that certification is, at best, a very crude indicator of teacher quality. At the very least, certification status is capable of differentiating the most ineffective teachers from all other teachers. Certification status at the early elementary level is not sensitive enough to distinguish quality past the standard level. The question is whether or not finer gradations of quality matter. Put another way, is being able to differentiate between the least effective teachers and all other teachers a rationale that justifies certification? In the context of value-added research, proponents argue that it does indeed.

According to value-added research by Sanders and Rivers (1996), the teacher effect is generally additive and cumulative, and not compensatory. They contend that even when students are assigned to a good teacher after successive poorer ones, the teacher effect is virtually lost. Simply, no number of good teachers can compensate for poorer ones, particularly in the early grades. These findings suggest that teachers with emergency certification are less effective than teachers with standard or (to a lesser extent) advanced certification at raising both reading and mathematics achievement. Given the reported importance of the literacy and numeracy skills gained in the early grades for future academic work, good teaching and teachers are critical at this level.

Therefore, limiting or prohibiting the use of emergency certified teachers at this level, especially for students at risk, could be a key strategy for ensuring long-term reading and mathematics achievement. Given that teacher shortages do not often materialize in the early grades, except with specialized populations such as English language learners and students with disabilities, it seems like a plausible strategy (Ingersoll, 2003). Coupled with ensuring that teachers are highly qualified (i.e. possessing at least standard certification, according to these data) at this grade level, attending to the instructional practices utilized by teachers seems to matter as well. The instructional effects are discussed below.

Emergency certified teachers utilize instructional practices that have a negative effect on student achievement in mathematics. In the arguments against certification, opponents could describe the effects of instructional practices as stronger than those of certification status in both reading and mathematics. But these findings indicate that the effects of certification status interact with the effects of teaching practices. Despite the

fact that no large differences were found in reported instructional practices by certification status, emergency certified teachers did report greater usage of number sense instruction and student-centered mathematics instruction, and each of these practices is negatively associated with student achievement. Moreover, when considering the interaction effect, the negative effect of number sense instruction is even greater when used by teachers with emergency certification than when used by teachers with standard certification.

In short, emergency certified teachers who use number sense instruction lower mathematics achievement. By itself, number sense instruction has a relatively small, negative effect—about 5% of an SD—on mathematics achievement. When emergency certification status is considered as part of the interaction, the effect size, while still small, increases by 200%. Emergency certification status thus exacerbates the negative effect of number sense instruction on end-of-first-grade mathematics achievement. When emergency certified teachers utilize it, mathematics achievement decreases by 15% of an SD. When the indirect effects of certification status are considered along with the effects of number sense instruction, the importance of both instructional practices and certification status is magnified—further advancing the argument for certification. In cases in which school systems need to continue to use emergency certified teachers, attending to their curricular coverage and instructional practices may be especially critical to students' learning.

Summary

Does certification matter? Certification opponents could contend that it does not, given the lack of direct effects on student achievement, the stronger effects of instructional practices, and the diminishing (or even detrimental) effects of earning higher levels of certification on student achievement. Proponents, however, could argue that certification matters because it can be used to discriminate between the best and the worst teachers—those who might widen the achievement gap and use instructional practices that limit student achievement. Although both arguments are convincing, the proponents' argument exposes potential inequities that might not be rectified without certification—chiefly, aspects of certification seem to matter, especially for low-achieving students. Moreover, students in the early grades are in a very unique position. In this era of standards-based accountability, getting students to learn (and demonstrate their learning in high-stakes testing in later grades) needs to start happening in the first grade if not before. In light of the value-added research, the inequities caused by an unequal distribution of good teachers could be mitigated to some extent by ensuring that the teachers of young children have well-qualified teachers; based on the findings from this study, that means teachers with at least standard certification who make use of instructional practices related to student achievement.

Though the case for equity is very important, the lesson to learn from certification opponents is that the current theory of action underlying the practice of certifying teachers (at least in the case of state-based licensure schemes) is not adequate. The theory of action put forth by credentialing regimes is that teachers' qualifications are tantamount

to quality and are conceivably operationalized through their practice. This theory of action confuses qualifications and quality. Shifting certification away from simply *what teachers know* toward *what teachers know, do, and do consistently with the hardest to teach students* will likely result in better policy and practice. Additionally, considering the effects of advanced certification poses another challenge to the current certification paradigm. Perhaps, for advanced certification to matter, it should mirror the rigor involved in obtaining National Board Certification, which is more closely linked to performance than simply course counting and time on the job.

In short, these data indicate that certification, while imperfect, does seem to matter. The goal for policymakers and educators will be to learn from the limitations of the current certification method and make substantial changes that increase the utility and efficacy of certification status as an indicator of teacher quality. The next section contextualizes these findings in the literature presented earlier.

Revisiting the Literature

At the heart of this study has been an attempt to further inform the literature on teacher quality generally and teacher certification particularly. Rice (2003) in her review of the empirical literature related to teacher quality boldly states that “Teacher quality matters. In fact, it is the most important school-related factor influencing student achievement” (p. v). In each of the teacher attributes she studies—teacher experience, teacher preparation programs, teacher coursework, teachers’ own test scores, and teacher certification—the literature, overall, is mixed, especially with respect to teacher certification. The literature reviewed in Chapter 2 with respect to teacher certification,

identified the lack of research concerning the effects of certification status at the elementary level as a significant gap in the record—a gap that this study can help to close. This section revisits three of the major bodies of literature reviewed in Chapter 2 to contextualize the findings of the present study, present broad implications of this work for educators, teacher educators, and policymakers, and poses future areas for research. These bodies include (a) the effects of teacher certification on student achievement; (b) the effects of certification and classroom instructional practices; and (c) empirical approaches for measuring the classroom instructional practices. The latter two are discussed together below.

Teacher Certification and Student Achievement

To review, the literature concerning the effects of teacher certification on student achievement focused primarily on high school teachers and students, where other grade levels were considered the sample sizes were relatively small. Findings concerning the effects of certification were either mixed or inconclusive despite the level of analysis. On balance, research demonstrated a positive effect for high school mathematics achievement when certified teachers had subject specific certification in mathematics⁷. In terms of type of certification—emergency, advanced, and alternative—the record is even

⁷It may be the case that “double certification” matters as Goldhaber and Brewer (2002) found with respect to already certified teachers with additional content area certification—although this pattern only held for mathematics. Conversely, the researchers found the opposite effect for certified teachers with English certification. Goldhaber and Brewer’s (2002) work begs the question of whether coupling general elementary certification with early childhood certification or subject-specific certification, for example, will matter for the achievement of young students. Future research may want to test this hypothesis, specifically. Hypothetically, given what is known about the relatively positive effects of additional pedagogical and content coursework conceivably undertaken by those earning such certification (see Monk, 1994; Monk & King, 1994) an effect may likely be present.

less straightforward. The evidence does not indicate any clear pattern of impact with respect to type of certification and student outcomes compared to teachers with standard certification.

The major contribution this study offers to the existing literature on teacher quality is twofold. First this study helps to fill the gap on the effects of certification status on the achievement of young students which uses a robust, nationally representative sample of public first grade teachers and their students. Second, the study provides insights into teachers' instructional practices which are of increasing interest with respect to teacher quality (see Wenglinsky, 2002). Overall, the findings of this study fit within the pattern of similar studies in which certification status—emergency, standard, and advanced—seem to have little or no direct effects on student achievement. The indirect effects associated with certification status and prior achievement indicate that when compared to standard certified teachers, emergency certified teachers may not be able to close the gap in achievement in reading and mathematics; while advanced certified teachers are unable to close the gap in mathematics. In terms of the examined instructional practices in reading and mathematics the findings suggest that practice did not vary by certification status but may vary in effectiveness for mathematics (e.g., number sense).

What accounts for the relatively unimpressive findings concerning the effects of certification status on student achievement in this study and others? The answers to this question concerns both the methodological issues involved with studying certification status using large-scale survey data and phenomenon itself.

In studying independent variables like certification status, researchers must grapple with the fact that these variables are largely defined by institutions, not by nature. As indicated in Chapter 1, certification status is a varied treatment. Requirements for certification vary from state to state and certification labels—such as emergency, provisional, standard, and advanced—are defined in any number of ways. For example, the requirements for obtaining advanced certification are substantively different in New York and South Carolina, yet both teachers from both states can report that they hold advanced certification. These construct variations make it difficult for researchers to measure certification consistently and reliably across research contexts. Yet the variability is inherent to phenomenon and the governing bodies that fine tune certification to local conditions and regulate teaching.

For researchers the variability is less fortuitous, especially in large-scale survey research in which teachers self-report their status based on their own states' certification definitions and not the definitions intended by the researchers. Two studies, using NELS:88 data highlight this dilemma. Goldhaber and Brewer (2000) reported that students of emergency certified teachers performed similarly in mathematics and science, as did their peers whose teachers held standard certification. In light of their findings, they suggest certification should be abandoned. In a rebuttal, Darling-Hammond, Berry, and Thoreson (2001) argued that some of the emergency certified teachers in Goldhaber and Brewer's sample were most likely experienced teachers who held some sort of licensure but were not fully certified in the state where they were currently teaching. These researchers conclude with calls for additional inquiry into how certification and teacher education operate. These two studies highlight the significance of certification

labels in general and the “emergency certified” label particularly. Additionally, the interpretations and recommendations lodged on their respective findings are significantly different.

In the present study, the direct effects of emergency certification in the unadjusted regression models and the indirect effects of emergency certification and prior reading and mathematics achievement are noteworthy, but are suspect to the same construct ambiguity discussed above. Are these teachers, in fact, already highly certified teachers in limbo, or are they the individuals who have “walked off the street” into teaching positions? There is no way to determine either definitely; however, the descriptive data seems to suggest that they are not Darling-Hammond, Berry, and Thoreson’s “highly certified” teachers. Like other studies which consider certification *prima facie*, the results are like to be unimpressive. The opportunity is to create better survey instruments which attempt to mitigate some of this variation. At the very least, conducting a construct validity study of the ECLS-K certification with a small sub-sample of the respondents may be helpful in clarifying the reliability of the type of certification status measure used in most NCES surveys. Future surveys may want to consider how to develop questions about certification status that capture more nuance and variability apparent in the phenomenon itself.

Certification Status and Teachers’ Instructional Practices

The notion of teacher quality explored in this paper is concerned with not only how effective teachers are by virtue of their certification status but also by how effective teachers are by virtue of what they do in the classroom. To review, the literature

reviewed in Chapter 2 concerning teacher certification status and instructional practices is emergent and inconclusive. The available evidence suggests that kindergarten teachers who did not hold ECE certification or elementary certification used less student-centered instructional practices. Standard and advanced teachers tended to emphasize higher-order thinking skills which deepened students' achievement in reading comprehension. There was no consistent evidence that reform-oriented instructional practices promoted gains in early grades in reading or mathematics achievement. The bulk of the research reviewed, as well as this study, is limited by data collection methods which rely on teacher self-report data to construct measures for instructional practices.

Despite these limitations, this study provides an alternative approach for investigating teacher quality by identifying instructional practice factors that might be used in future research. More importantly, the study suggests that it is not only the use of practices that should be examined but the effectiveness of the practices when performed by teachers with different levels of preparation and experience. It may also be important to consider whether such practices are equally effective for all groups of students. In other words, what constitutes quality teaching may depend on not only who engages in the practices but who the students are in the classroom.

But what accounts for the modest variability between teachers with different certification status in instructional practices found in this study? The mostly likely reason is how instructional practices are measured. Despite the statistical strength of the examined instruction practice composites, a more fundamental problem exists in the way in which classroom practice data are collected using surveys. Rowan et al. (2002) summarize this problem by stating:

A great deal of research on the ways in which respondents complete questionnaires suggests that the kinds of questions asked on [large-scale survey] teacher questionnaires—questions about how much time was spent in routine forms of instructional activities—cannot be responded to accurately in “one-shot” questionnaires. This lack of accuracy probably introduces substantial error into our analyses, biasing all effect sizes downward and perhaps preventing us from discovering statistically significant relationships among teaching processes and student achievement.

The presence of measurement error attenuates relationships with all of the other variables in a model, making it more difficult, as Rowan and his colleagues observe, to determine differences in instructional processes between groups of teachers and the actual effects of these processes on student learning.

Clearly better measurement of instructional practices is both necessary and desirable. The solution, however, may not be as simple as getting “closer to the action” as some (Mayer, 1999; Stecher et al, 2002,) have suggested. Studies which involve participant observations and videotape studies could potentially offer a more nuanced and much needed view of instructional practices. These data collection methods are not a panacea, though. In fact, coming to consensus on what was viewed, when it was viewed, and the context for practice present formidable challenges for such investigations. Even with extensive norming for observers, it is likely that measurement error will still limit our ability to use practice routinely and consistently as an indicator of teacher quality without consensus on what constitutes a teacher’s practice.

Underneath these methodological concerns, may be a more fundamental problem with studying teachers' instructional practice—namely, claims that teacher's lack a shared professional knowledge-base of teaching (Hiebert, Galimore, & Stigler, 2002). The utility of a professional knowledge-base is twofold. Professional knowledge is prescriptive in the sense that practitioners know what they are expected to do and diagnostic in the sense that practitioners are able to predict outcomes and problematize practice. Without a shared knowledge base, both researchers and policy makers may find it difficult to identify exactly what teachers do in the classroom that makes a difference. Although teacher practices may be an important alternative to certification in understanding teacher quality, it is not without its own inherent ambiguities and epistemological problems.

In sum, the focus on elementary grades remains relatively unstudied with respect to teacher certification and the independent variables explored in this study. The findings from this study contribute to both the existing record on different levels of teacher certification and their effects on student achievement in the early grades as well as the existing record on instructional practices and their effects on student achievement in the early grades. The intersection of these findings suggest that if certification status matters, it may matter in subtle ways that reflect the effectiveness of practices and the ability of teachers to boost the achievement of students who enter the early grades with less knowledge in reading and mathematics. Nonetheless, the results from this study and related studies reveal that there remain significant methodological hurdles to overcome before we fully understand the effects of certification status and instructional practices on student achievement in elementary schools. Clearly more research in this area may help

such a translation come to fruition. The next section identifies some limitations of this study and recommendations for future research.

Limitations and Recommendations for Future Research

The current study has several limitations. First, the study employed analyses for which it was necessary to average student data for each teacher rather than to perform the analysis at the student level. Other analytic options (e.g., hierarchical linear modeling) were explored but found to be problematic given the characteristics of the current data. Specifically, a large proportion of the teachers had a very small number of students (or even a single student) included in the data set. Regardless of the methodological technique employed, because the primary independent variable was a teacher-level variable, the averaging of student data within each classroom was necessary. If these estimates of classroom achievement are unreliable, they may underestimate the true effects of teacher certification and practices in the models.

Although the ECLS-K dataset provides many advantages, including its longitudinal design, the number of students associated with teachers in grades after kindergarten limit the analytic techniques than can be employed and raise the possibility of underestimating teacher effects on student learning. This problem has long plagued the general-purpose surveys conducted by the National Center for Education Statistics (NCES), including High School and Beyond and the National Educational Longitudinal Study. It would be helpful if researchers and policymakers worked with local, state, and national agencies to design studies that permit more reliable and fine-grained indicators

of certification status, teacher practices and other teacher effects on the achievement of students in their classes.

Second, the primary independent variable in this study, certification status, as discussed earlier is problematic. ECLS-K constructed responses to the certification status in such away that emergency certification is also included with “temporary” and “probationary” classifications. In practice, temporary and provisional status can be more similar to regular and standard certification than it is to emergency certification. Furthermore, it was not possible to disaggregate the original response category; therefore the constructed emergency certification category used in this analysis may be biased. Nonetheless, the descriptives show that the teachers in the constructed emergency category are similar in profile to other emergency teachers reported in the literature.

Despite this limitation, it would be of interest to replicate the current study using the disaggregated approach described in the conceptual framework section. Under this approach the inquiry would focus on the separate and relative effects of the criteria—degrees, courses, and experience—which constitute certification. Moreover, this analysis focused on high-level certification status. As discussed earlier, many types of certifications are available to elementary teachers. For example, these teachers can earn elementary certification, early childhood education certification, special education certification, or a combination of each. A subsequent analysis could examine different configurations of certification and explore their relationship outcome variables associated with teacher quality, specifically student achievement and instructional practices.

Finally, the validity of identified instructional practice variables is limited to the extent to which teachers reported their classroom activities accurately. Many researchers

acknowledge that teachers misreport, for a variety of reasons, their classroom practices. The sources of the misreports are not simply dubious cases of what a teacher believes she should report but does not actually do; rather, they are more likely due to the difficulties associated with measuring teachers' classroom practices generally. For example, there can be wide range of interpretation concerning a classroom practice like "cooperative learning." Although the use of self-reported practices poses a challenge to the validity of any study, survey data remains a fundamental, albeit potentially problematic strategy for exploring the possible effects of teacher practices on student learning. Future researchers may want to further investigate the instructional practice measures developed in this analysis and others through a mixed-methods study.

Conclusion

This study sought to identify the relationship between certification status, a teacher's instructional practices, and first grade reading and mathematics achievement. The findings from this study stop short of recommending that certification be rejected out of hand, but they do caution policymakers against using certification as a reliable indicator of teacher quality or a strategy for engendering teacher quality. Interestingly, little variation existed between the three categories of teachers with respect to their classroom practices, though there was an indication that the effects of at least number sense varies with certification status.

Good teachers matter, and they matter most for the students who need them the most. Ensuring a high-quality teacher workforce is critically important to ensuring equitable outcomes for students. This study highlights potential differences between

creating highly qualified teachers and high-quality teachers. The distinctions between the two are important and may mean the difference between promising equitable outcomes for students and actually realizing them.

APPENDIX A: THE NATIONAL CENTER FOR ALTERNATIVE CERTIFICATION'S
CATEGORIZATION OF ALTERNATE ROUTE AND CERTIFICATION PROGRAMS

Category Title	Description
CLASS A	<p>This category is reserved for programs that meet the following criteria:</p> <ul style="list-style-type: none"> • The alternative teacher certification route has been designed for the explicit purpose of attracting into elementary and secondary school teaching talented individuals who already have at least a bachelor's degree in a field other than education. • The alternate route is not restricted to shortages, secondary grade levels, or subject areas. • These alternative teacher certification routes involve teaching with a trained mentor, and formal instruction that deals with the theory and practice of teaching during the school year—and sometimes in the summer before and/or after.
CLASS B	<p>These are teacher certification routes that have been designed specifically to bring into teaching talented individuals who already have at least a bachelor's degree. These routes involve specially designed mentoring and formal instruction. However, states in this class restrict the program to shortages, secondary grade levels, and/or subject areas.</p>
CLASS C	<p>These routes entail review of academic and professional background, and transcript analysis. They involve specially (individually) designed in-service and course-taking necessary to reach competencies required for certification, if applicable. The state and/or local school district have major responsibility for program design.</p>
CLASS D	<p>These routes entail review of academic and professional background, and transcript analysis. They involve specially (individually) designed in-service and course-taking necessary to reach competencies required for certification, if applicable. An institution of higher education has major responsibility for program design.</p>
CLASS E	<p>These postbaccalaureate programs are based at an institution of higher education.</p>

CLASS F	These programs are basically emergency routes. The prospective teacher is issued some type of emergency certificate or waiver that allows the individual to teach, usually without any on-site support or supervision, while taking the traditional teacher education courses requisite for full certification.
CLASS G	Programs in this class are for persons who have few requirements left to fulfill before becoming certified through the traditional approved college teacher education program route—for example, persons certified in one state moving to another, or persons certified in one endorsement area seeking to become certified in another.
CLASS H	This class includes those routes that enable a person who has some “special” qualifications, such as a well-known author or Nobel Prize winner, to teach certain subjects.
CLASS I	These states reported that they were not implementing alternatives to the approved college teacher education program route for licensing teachers.
CLASS J	These programs are designed to eliminate emergency routes. They prepare individuals who do not meet basic requirements to become qualified to enter an alternate route or a traditional route to teacher licensing.
CLASS K	These avenues to certification accommodate specific populations for teaching, such as Teach for America, Troops to Teachers, and college professors who want to teach in K–12 schools.

Source: Feistritz, E. (2004). *Alternative teacher certification: A state-by-state analysis 2004*. Washington, DC: National Center for Education Information.

APPENDIX B: SUMMARIES OF SELECTED RESEARCH GROUPED BY TOPIC

Table B1. Research Comparing Teacher Certification Status and Student Achievement

Citation	Findings	Method/Sample
<p>Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. <i>Educational Evaluation and Policy Analysis</i>, 23, 57–77.</p>	<p>D-H disputes Goldhaber & Brewer’s (2000) findings that teachers with emergency certification are as good as those with standard certification. She suggests that the emergency label is a misnomer and that the sample of emergency certified teachers studied resembles teachers with standard certification. Goldhaber & Brewer (2001) issued a rejoinder to D-H’s claims.</p>	<p>NELS:88</p>
<p>Darling-Hammond, L. (1999). <i>Teacher quality and student achievement: A review of state policy evidence</i>. Seattle: University of Washington, Center for the Study of Teaching Policy.</p>	<p>Highly publicized policy report in which the researcher asserts that the strongest correlates to student achievement in reading and mathematics are certification and teacher preparation. The results have been challenged by others as suffering from “aggregation bias.” Aggregation bias, Hanushek (1996) argues, causes different interpretations of data at the student, teacher, and school level.</p>	<p>State-level data using NAEP results; controls for poverty and language status</p>

Citation	Findings	Method/Sample
<p>Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. <i>Educational Evaluation and Policy Analysis</i>, 22, 129–145.</p>	<p>Researchers looked how a high school teacher’s certification status predicts student achievement in high school mathematics. They determined that students whose teachers held standard certification did better than their peers whose teachers held private school certification or were not certified in the subject area. However, they found that the students whose teachers held emergency certification did no worse than students whose teachers had standard certification.</p>	<p>Econometric analysis; NELS:88</p>
<p>Goldhaber, D., & Brewer, D. (1998). Why should we reward degrees for teachers? <i>Phi Delta Kappan</i>, 10, 134–138.</p>	<p>Authors suggest that teacher certification, teacher experience, and possessing a master’s degree were not related to higher test scores for 10th grade students. Although subject matter coursework had a small but significant effect on student test scores, certification in English had a statistically significant negative effect.</p>	<p>NELS:88; controlled for prior achievement; also linked student records to teachers</p>

Citation	Findings	Method/Sample
<p>Laczko-Kerr, I., & Berliner, D. C. (2002). The effectiveness of “Teach for America” and other under-certified teachers on student academic achievement: A case of harmful public policy [Electronic version]. <i>Education Policy Analysis Archives</i>, 10. Retrieved September 3, 2003, from http://epaa.asu.edu/epaa/v10n37/</p>	<p>The researchers determine that fully certified teachers are more effective than undercertified teachers. Undercertified teachers are those who have provisional, emergency, or temporary status. Though the literature generally defines emergency certification, no operational definition of <i>undercertified</i> was given to compare how undercertified teachers’ qualifications differed from those of certified teachers. In addition, the authors’ preoccupation with disproving the merits of TFA was readily apparent, and political motives for doing so seemed less than transparent.</p>	<p>Match-pairing design of 293 teachers in Arizona; SAT-9 (language, reading, and mathematics scores were used); no controls (prior student achievement, or SES) were used in this analysis</p>
<p>National Center for Education Statistics. (1992). <i>1990 science report card</i>. Washington, DC: U.S. Department of Education.</p>	<p>The report states that none of the following characteristics are statistically associated with student achievement: teacher’s experience, certification type or level, master’s degree, coursework in the subject matter. These findings are suspicious because of the aggregation bias problem.</p>	<p>NAEP data, aggregated to national level</p>
<p>Qu, Y., & Becker, B. (2003, April). <i>Does traditional teacher certification imply quality?</i> Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.</p>	<p>The researchers undertake a meta-analysis of studies related to the differences between alternatively certified teachers, traditionally certified teachers, and emergency certified teachers. They find that traditionally certified teachers and alternatively certified teachers “perform equivalently.” Traditionally certified teachers outperformed emergency certified teachers. Comparisons of out-of-field teaching were mixed.</p>	<p>Meta-analysis (synthesis of research); uses 24 studies; selection criteria are apparent</p>

Table B2. Studies Comparing Alternatively Certified Teachers With Traditionally Certified Teachers

Citation	Findings	Method/Sample
Ashton, P. (1996). Improving the preparation of teachers. <i>Educational Researcher</i> , 25, 21–22.	Researchers reported that certified teachers receive higher supervisor ratings than uncertified teachers.	Review of Research
Brown, D., Edington, E., Spencer, D. A., & Tinafero, J. (1989). A comparison of alternative certification, traditionally trained, and emergency permit teachers. <i>Teacher Education and Practice</i> , 5, 21–23.	Study investigated the viability of an alternative certification program at the University of Texas, El Paso. Traditionally certified teachers were compared to alternatively certified teachers and emergency teachers. Brown found no differences between teachers on several outcomes. However, when compared with alternate route teachers, emergency permit teachers did better on all but one outcome.	63 Texas teachers, ANOVA
Decker, P., Mayer, D., & Glazerman, S. (2004). <i>The effects of “Teach for America” on students: Findings from a national evaluation</i> . Princeton, NJ: Mathematica Policy Research.	The researchers compared the Teach for America (TFA) teachers to other teachers, new and veteran, in similar types of schools and determined that students of TFA teachers made stronger gains in mathematics than did students of all other teachers, and did as well as students of all other teachers in reading. “Even though Teach For America teachers generally lack any formal teacher training beyond that provided by Teach For America, they produce higher test scores than the other teachers in their schools—not just other novice teachers or uncertified teachers, but also veterans and certified teachers” (p. 3).	17 high-poverty schools in 6 regions around the country; students randomly assigned to TFA or non-TFA teachers; standardized test (Iowa Test of Basic Skills)

Citation	Findings	Method/Sample
<p>Guyton, E., Fox, M., & Sisk, K. (1991). Comparisons of teaching attitudes, teacher efficacy, and teacher performance of first-year teachers prepared by alternative and traditional teacher education programs. <i>Action in Teacher Education</i>, 13(2), 1–9.</p>	<p>Compared two groups of teachers: alternatively certified and traditionally certified. The groups were similar in terms of gender, subject matter taught, and student SES. The teachers were similar on almost all measures.</p>	<p>Sample included 23 alternatively certified teachers and 26 traditionally certified teachers; Regression</p>
<p>Lutz, F., & Hutton, J. (1989). Alternative teacher certification: Its policy implications for classroom and personnel practice. <i>Education Evaluation and Policy Analysis</i>, 11(3), 237–254.</p>	<p>Researchers found that principals and mentor teachers, across all grade levels, rated alternatively certified teachers as high or higher than first-year teachers with standard (i.e. initial) certification.</p>	<p>100 alternatively certified teachers in Texas</p>
<p>Miller, J., McKenna, M., & McKenna, B. (1998). A comparison of alternatively and traditionally prepared teachers. <i>Journal of Teacher Education</i>, 49, 165–176.</p>	<p>Researchers found no differences in the student achievement, instructional practices, or perceptions of teaching ability of alternatively certified teachers.</p>	<p>Paired comparison analysis; 41 alternatively certified and 41 traditionally certified 5th and 6th grade teachers</p>
<p>Raymond, M., Fletcher, S. H., & Luque, J. (2001). <i>Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas</i>. Unpublished manuscript, Stanford University, Stanford, CA.</p>	<p>The researchers determined that being taught by TFA teachers was generally positive; the average TFA teacher and average non-TFA teacher performed similarly and the differences between the two groups were not statistically significant.</p>	<p>TFA and non-TFA teachers and their students in Houston Independent School District; TAAS test score data</p>

Table B3. Studies Examining Certification Based on Teacher Grade Level or Subject Area

Citation	Findings	Method/Sample
<p>Felter, M. (1999). High school staff characteristics and mathematics achievement. <i>Education Policy Analysis Archives</i>, 7(9). Retrieved April 5, 2004, from http://epaa.asu.edu/epaa/v7n9.html</p>	<p>Felter finds that students of teachers with full certification scored higher on mathematics exams. Analysis is conducted at school level.</p>	<p>Correlational analysis and multivariate regression controlling for poverty. 797 California high schools.</p>
<p>Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. <i>Educational Evaluation and Policy Analysis</i>, 22, 129–145.</p>	<p>Researchers looked how a high school teacher’s certification status predicts student achievement in high school mathematics. They determined that students whose teachers held standard certification did better than their peers whose teachers held private school certification or were not certified in the subject area. However, they found that students whose teachers held emergency certification did no worse than students whose teachers had standard certification.</p>	<p>Econometric analysis; NELS:88</p>
<p>Goldhaber, D., & Brewer, D. (1997). Evaluating the effect of teacher degree level on educational performance. In W. Folwer (Ed.), <i>Developments in school finance</i>. Washington, DC: U.S. Department of Education.</p>	<p>Researchers determined that certification has a small effect on student achievement in mathematics and science.</p>	<p>Econometric model utilizing OLS regression using, NELS:88 student sample (n=18,000)</p>

Citation	Findings	Method/Sample
<p>Goldhaber, D., & Brewer, D. (1996). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. <i>Journal of Human Resources</i>, 32, 505–523.</p>	<p>Students whose secondary school mathematics teachers were certified in mathematics or who had majored in mathematics or earned a master's degree in mathematics had higher scores than their peers whose teachers did not.</p>	<p>Multivariate analysis; NELS:88</p>
<p>Hawk, P., Coble, C., & Swanson, M. (1985). Certification: It does matter. <i>Journal of Teacher Education</i>, 36, 13–15.</p>	<p>The students of secondary school teachers who were certified in mathematics scored higher than those of teachers not certified in mathematics.</p>	<p>Paired analysis; middle school and high school teachers; 36 teacher, 826 students</p>
<p>Mandeville, G. K., & Liu, Q. (1997). The effect of teacher certification and task level on mathematics achievement. <i>Teaching and Teacher Education</i>, 13, 397–407.</p>	<p>In this study, researchers compared the achievement scores of 7th grade students whose teachers held either secondary or elementary certification. Holding secondary certification was a proxy for “high,” or intensive, mathematics preparation, while “low,” or minimal, preparation was operationalized as holding elementary certification. Researchers found that for students whose teachers had the highest level of preparation (i.e. advanced certification in mathematics), this only slightly improved their scores on three measures of mathematics achievement.</p>	<p>Match-paired design; 9,000 7th grade students, 33 match pairs of schools whose teachers differed in mathematics subject knowledge preparation; controls for SES, urbanicity, school size, organization of school (6–8; 7–9); no control for prior achievement in mathematics</p>

Citation	Findings	Method/Sample
<p>Rowan, B., Correnti, R., & Miller, R. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects students of elementary schools. <i>Teachers College Record</i>, 104(8), 1525–1584.</p>	<p>For elementary teachers, subject-specific certification was not related to student achievement growth in mathematics or reading.</p>	<p>Hierarchical linear growth models; PROSPECTS</p>

Table B4. Empirical Research Focused on the Relationships Between Classroom Practices and Student Achievement

Citation	Findings	Method/Sample
<p>Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? <i>Educational Evaluation and Policy Analysis</i>, 21, 29–45.</p>	<p>The self-report surveys of instructional practices measured how frequently teachers engaged in 17 strategies related to algebra instruction. The researcher determined that the survey had construct validity after correlating observational data with self-report data. Drawbacks to the survey included that teachers' estimates of strategies used were unreliable and inconsistent in terms of a teacher's use of strategies could not be gauged by the survey.</p>	<p>Exploratory study; author-created survey instrument</p>
<p>National Center for Education Statistics. (1996). <i>High school seniors' instructional experiences in science and mathematics</i>. Washington, DC: U.S. Government Printing Office.</p>	<p>Focusing on higher-order thinking skills had a positive effect on 10th grade students' mathematics performance but not science performance.</p>	<p>NELS:88</p>

Citation	Findings	Method/Sample
<p>National Center for Education Statistics. (1999). <i>What happens in classrooms? Instructional practices in elementary and secondary schools, 1994–95</i> (NCES 1999-348). Washington, DC: U.S. Department of Education.</p>	<p>This descriptive study examined the impact of several teacher characteristics (experience, degrees attained, beliefs about student ability,) and student characteristics (English proficiency, income, and race) on classroom practices. A particular emphasis was placed on teachers who engaged in reform-oriented practices rather than conventional (or traditional) practices. Reform-oriented practices were defined as those emphasizing authentic assessment through portfolios, higher-order thinking skills, grouping strategies, and student-student talk. With respect to teacher characteristics, teachers who believed their students were more able (i.e., gifted) used the recommended strategies less frequently than teachers whose students were from a minority group. Teachers with more experience (more than 5 years) tended to use the recommended strategies more than teachers with less experience. Teachers with advanced degrees were more likely to use the recommended strategies than teachers with a four-year degree. Teachers who had participated in professional development on a recommended strategy the previous year were more likely to use the recommended strategies.</p>	<p>Descriptive analysis; 1994–1995 Teacher Follow-Up Survey; K-12 teachers</p>

Citation	Findings	Method/Sample
<p>Von Secker, C., & Lissitz, R. (1999). Estimating the impact of instructional practices on student achievement. <i>Journal of Research in Science Teaching</i>, 36, 1110–1126.</p>	<p>This study seeks to understand the effect of reform-oriented instructional practices (laboratory inquiry, critical thinking, and reducing teacher-centered instruction) for 10th grade high school science teachers and the effects on student achievement in science. Using an HLM model, the researchers find that the use of reform-oriented instructional practices was not associated with significant differences in the mean achievement of a school’s students. Furthermore, the authors note that as the frequency of reform-oriented practices increased, the disparity between the scores of minority and nonminority students and of female and male students decreased as well—raising for the authors serious equity concerns.</p>	<p>HLM; students (avg = 12) nested in schools ($n = 163$); 1990 High School Effectiveness Study (part of NELS:90); nationally representative sample of 10th graders</p>
<p>Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. <i>Education Policy Analysis Archives</i>, 10. Retrieved September 16, 2004, from http://epaa.asu.edu/epaa/v10n12/</p>	<p>The researcher finds that the effects of classroom practices were stronger than the effects of professional development and the use of higher-order thinking strategies when SES and class size are used as controls. When classroom practices are added to other teacher characteristics, the effects are similar in size to those of student background. The author concludes that teacher effects contribute just as much to student learning as student effects.</p>	<p>Multilevel structural equation modeling; 1996 NAEP mathematics data; 7,146 8th graders</p>

Table B5. School Effectiveness Studies and Early Empirical Studies on Classroom Instruction and Student Outcomes

Citation	Findings	Method/Sample
Brophy, J., & Good, T. (1986). Teacher-effects results. In M. C. Wiltrock (Ed.), <i>Handbook of research on teaching</i> (3rd ed.). New York: Macmillan Press.	Summarizes the literature on teacher effects as part of the school effectiveness movement.	Review and synthesis of related literature
Cooley, W. W., & Leinhardt, G. (1978). The instructional dimensions study. <i>Educational Evaluation and Policy Analysis</i> , 2, 7–25.	The authors studied the relationship between four classroom processes (opportunity, motivators, instructional events, and structure) and the effect of each on the achievement of low-income children in grades 1, 2, and 3. The authors conclude that opportunity showed the strongest positive correlation to achievement. Opportunity is defined as testing what is actually taught. The authors conclude that <i>what</i> is taught is a better predictor of student achievement than <i>how</i> something is taught.	400 low-income children in grades 1–3
Flanders, N. (1960). <i>Teacher influence, pupil attitudes and achievement</i> . Minneapolis: University of Minnesota.	Early study that helped initiate quantitative approaches to understanding teacher effects on student achievement. The focus was on classroom instruction.	Correlation; other associational measures
Medley, D., & Mitzel, H. (1959). Some behavioral correlates of teacher effectiveness. <i>Journal of Educational Psychology</i> , 50, 239–246.	Early study of 49 first-year NYC elementary school teachers in which the researchers measured various teacher behaviors (instructional practices) and how these practices correlated to student achievement.	49 first-year NYC teachers; differences in means between teachers

Citation	Findings	Method/Sample
<p>Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. <i>Elementary School Journal</i>, 83, 427–452.</p>	<p>This study reviews several studies (of varying types) of school effectiveness research. The research indicates that effective schools are those that have structure and order, purposefulness, and a human atmosphere, and use appropriate instructional techniques. The authors conclude that while the attributes of effective schools have been identified to have varying degrees of association with student achievement, little has been written about how to translate the research into practice.</p>	<p>Review of research</p>

Table B6. Studies Focused on the Instructional Practices of Teachers of Young Children

Citation	Findings	Method/Sample
<p>Berends, M., Chun, J., Schuyler, G., Stockly, S., & Briggs, R. (2002). Challenges of conflicting school reforms: Effects of new American schools in a high-poverty district. Santa Monica, CA: RAND. Retrieved August 19, 2004, from http://www.rand.org/publications/MR/MR1483/</p>	<p>This study examines the implementation of (National Academies of Science) NAS reforms in a large urban district in Texas. The researchers use an MLM to analyze the impact of (reformlike) instructional conditions—including teacher-reported collaboration, quality of professional development, and reformlike instructional practices—and achievement in reading and mathematics of 4th grade students. Using controls for SES, prior achievement, and race, the findings indicate that pedagogical decisions were not related to increased achievement in mathematics or reading on two measures of achievement (SAT-9 and TAAS). The authors explain that the lack of a relationship may be due to several issues, chief among them that the study was conducted when teachers were not familiar with the reform strategies.</p>	<p>4th grade TAAS scores in reading and mathematics; 3,800 students in 280 classrooms at 64 elementary schools; multilevel linear modeling</p>

Citation	Findings	Method/Sample
<p>Cohen, D., & Hill, H. (2000). Instructional policy and classroom performance: The mathematics reform in California. <i>Teachers College Record</i>, 102, 294–343.</p>	<p>Researchers determine that a teacher’s practice in mathematics is strongly influenced by policy measures that shape instruction, curriculum, and assessments. Using an OLS model, the researchers speculate that a teacher’s classroom practice bridges policy aims and student achievement. Teachers who understood the instructional policy that emphasized standards-based instruction and received professional development on it used more practices consistent with the reform rather than conventional practices.</p>	<p>OLS; California Learning Assessment System (CLAS); teacher survey instrument given to 4th grade teachers</p>
<p>Klein, S., Hamilton, L., McCaffrey, D., Stecher, B., Robyn, A., & Burroughs, D. (2000). <i>Teaching practices and student achievement: Report of first-year findings from the “Mosaic Study” of systemic initiatives in mathematics and science</i> (MR-1233-EDU). Santa Monica, CA: RAND.</p>	<p>Researchers examine first-year findings from a study of reform-oriented curricular implementations in mathematics and science to determine whether there is a relationship between student achievement and teachers’ use of reform-oriented practices (cooperative learning groups, inquiry-based activities, use of materials and manipulatives, and open-ended assessment techniques). Klein et al. find a weak relationship between a teacher’s instructional practice and student achievement in mathematics and science. Furthermore, there was no observed difference in the frequencies of teachers’ use of reform or traditional practice. Klein et al. explain the weak effects as due to methodological issues involved in large-scale research studies.</p>	<p>6 sites around the country using National Science Foundation–funded mathematics/science curricula; used controls of student background</p>

Citation	Findings	Method/Sample
<p>McMillian, J. (2003). The relationship between instructional and classroom assessment practices of elementary teachers and student scores on high-stakes tests (Report No. TM034718). Virginia: ERIC Clearinghouse on Assessment and Evaluation. (ERIC Document Reproduction Service No. ED472164) Retrieved on December 17, 2004 from http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/00000000b/80/28/04/d0.pdf</p>	<p>The researcher examined the relationship between instructional and classroom assessment (i.e. teacher-made assessments) and student achievement on a high-stakes test in reading and mathematics. The results of the analysis revealed that instruction and assessment are related to student achievement in reading and mathematics. Though the effects are small, cooperative learning, direct instruction, formative assessments, and essay tests were positively associated with achievement.</p>	<p>79 teachers; Multivariate regression analysis</p>
<p>Stecher, B., Hamilton, L., Ryan, G., Vi-Nhuan L., Williams, V., Robyn, A., & Alonzo, A. (2002). <i>Measuring reform-oriented instructional practices in mathematics and science</i> (DRU-2787-EDU). Santa Monica, CA: RAND Corporation.</p>	<p>This is a companion study to Klein et al. (2000). Mosaic II attempts to construct multiple measures of “reform-oriented” instructional practices based on surveys, responses to scenarios, teacher logs, teacher interviews, and observational instruments.</p>	<p>Instrument construction</p>
<p>Vartuli, S. (1999). How early childhood teacher beliefs vary across grade level. <i>Early Childhood Research Quarterly</i>, 14, 489–514.</p>	<p>The researcher determined that as grade level increased, teachers’ self-reports of classroom practices correlated with developmentally appropriate practices (DAP) decreased. Less experienced teachers and those with early childhood education certification were more likely to believe in and use DAP.</p>	<p>Early Childhood Beliefs and Practices Survey (Marcon, 1988); Teacher Beliefs Scale (Charlesworth,. 1998)</p>

REFERENCES

- Ackerman, D. (2004). States' efforts in improving qualifications of early care and education teachers. *Educational Policy*, 18, 311–337.
- American Board for the Certification of Teacher Excellence. (2004). www.abcte.org
- Anderson, J. (1995). Who's in charge? Teachers' views on control over school policy and classroom practices. *Office of Educational Research and Improvement Research Reports*. Washington, DC: U.S. Department of Education, Office of Education, Research, and Improvement.
- Ashton, P. (1996). Improving the preparation of teachers. *Educational Researcher*, 25, 21–22.
- Bacharach, S. (1990). *Education reform: Making sense of it all*. Boston: Allyn and Bacon.
- Ballou, D., & Podgursky, M. (2000). Reforming teacher preparation and licensing. *Teachers College Record*, 102(1), 5–27.
- Berends, M., Chun, J., Schuyler, G., Stockly, S., & Briggs, R. (2002). Challenges of conflicting school reforms: Effects of new American schools in a high-poverty district. Santa Monica, CA: RAND. Retrieved August 19, 2004, from <http://www.rand.org/publications/MR/MR1483/>
- Billingsley, B. S. (1993). Teacher retention and attrition in special and general education: A critical review of the literature. *The Journal of Special Education*, 27, 137–174.
- Bond, L., Jaeger, R., Smith, T., & Hattie, J. (2000). The certification system of the National Board of Professional Teaching Standards: A construct and consequential validity study. Center for Educational Research and Evaluation, University of North Carolina at Greensboro.
- Bowman, B., Donovan, S., & Burns, S. (Eds.). (2001). *Eager to learn: Educating our preschoolers*. Report of the National Research Council, Committee on Early Childhood Pedagogy, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.
- Brandt, R. (1993). On restructuring roles and relationships: A conversation with Phil Schlechty. *Educational Leadership*, 51(2), 8–11.

- Bredekamp, S., (Ed.). (1987). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8*. Washington, DC: National Association for the Education of Young Children.
- Bredekamp, S. (1993). The relationship between early childhood education and early childhood special education: Healthy marriage or family feud? *Topics in Early Childhood Special Education, 13*(3), 258–273.
- Brookover, W., & Lezotte, L. (1979). Changes in school characteristics coincident with changes in student achievement. East Lansing, MI: Institute for Research on Teaching.
- Brophy, J., & Good, T. (1986). Teacher-effects results. In M. C. Wiltrock (Ed.), *Handbook of research on teaching* (3rd ed.). New York: Macmillan Press.
- Brown, D. (1995). *Degrees of control: A sociology of educational expansion and occupational credentialism*. New York: Teachers College Press.
- Brown, D., Edington, E., Spencer, D. A., & Tinafero, J. (1989). A comparison of alternative certification, traditionally trained, and emergency permit teachers. *Teacher Education and Practice, 5*, 21–23.
- Brownell, M., & Smith, S. (1992). Attrition/retention of special education teachers: Critique of current research and recommendations for retention efforts. *Teacher Education and Special Education, 15*, 229–248.
- Burstein, L., McDonnell, L. M., Van Winkle, J., Ormseth, T., Mirocha, J., & Guiton, G. (1995). *Validating national curriculum indicators* (MR-658-NSF). Santa Monica, CA: RAND.
- Burts, D., Hart, C., Charlesworth, R., Fleege, P., Mosley, J., & Thomasson, R. (1992). Observed activities and stress behaviors of children in developmentally appropriate and inappropriate kindergarten classrooms. *Early Childhood Research Quarterly, 7*, 297–318.
- Charlesworth, R. (1998). Developmentally appropriate practice is for everyone. *Childhood Education, 74*(5), 274–282.
- Clandinin, D. J., & Connelly, F. M. (1987). Teachers' personal knowledge: What counts as personal studies of the personal. *Journal of Curriculum Studies, 19*, 487–500.
- Clandinin, D. J., & Connelly, F. M. (1995). *Teachers' professional knowledge landscapes*. New York: Teachers College Press.
- Clandinin, D. J., & Connelly, F. M. (1996). Research in progress: Studying teacher knowledge. *Alberta Journal of Educational Research, 52*(1), 67–69.

- Clifford, G. (1989). Man/woman/teacher: Gender, family, and career in American educational history. In Donald Warren (Ed.), *American teachers: Histories of a profession at work* (pp. 293–343). New York: Macmillan.
- Cohen, D., & Hill, H. (2001). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record*, *102*, 294–343.
- Cohen, D., Raudenbush, S., & Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*(2), 119–142.
- Coleman, J., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, DC: National Center for Education Statistics.
- Coles, G. (2003). *Reading the naked truth: Literacy, legislation, and lies*. New York: Heinemann.
- Council of Chief State School Officers. (1992). *Model standards for beginning teacher licensing and development: A resource for state dialogue*. Retrieved September 4, 2002, from <http://www.ccsso.org/content/pdfs/corestrd.pdf>
- Croninger, R., Rice, J., Rathbun, A., & Nishio, M. (in press). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review*.
- Darling-Hammond, L. (1990). Teaching and knowledge: Policy issues posed by alternative certification for teachers. *Peabody Journal of Education*, *67*(3), 123–154.
- Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence*. Seattle: University of Washington, Center for the Study of Teaching Policy.
- Darling-Hammond, L. (2000). Solving the dilemmas of teacher supply, demand, and standards: How we can ensure a competent, caring, and qualified teacher for every child. New York, NY: National Commission on Teaching and America's Future.
- Darling-Hammond, L., Berry, B., Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis*, *23*, 57–77.
- Darling-Hammond, L., Wise, A., & Klein, S. (1995/1999). *A license to teach: Raising standards for teaching*. San Francisco: Jossey-Bass.

- Decker, P., Mayer, D., & Glazerman, S. (2004). *The effects of "Teach for America" on students: Findings from a national evaluation*. Princeton, NJ: Mathematica Policy Research.
- Dee, T. (2003). Teachers, race, and student achievement in a randomized experiment. Unpublished manuscript.
- Denton, K., and West, J. (2002). Children's reading and mathematics achievement in kindergarten and first grade (NCES 2002-125). Washington, DC: National Center for Education Statistics.
- Diamond, J., & Spillane, J. (2002). *High stakes accountability in urban elementary schools: Challenging or reproducing inequality*. Evanston, IL: Institute for Policy Research. Retrieved December 5, 2003, from <http://www.northwestern.edu/ipr/publications/papers/2002/WP-02-22.pdf>
- Dingwall, R., & Fenn, P. (1987). A respectable profession? Sociological and economic perspectives on the regulation of professional services. *International Review of Law and Economics*, 7(1), 51-64.
- Dossett, D., & Munoz, M. (2000). Educational Reform in the accountability era: The impact of prior achievement and socio-economic conditions on academic performance (ERIC Document Reproduction Service No. ED 468491)
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15-24.
- Education Commission of the States. (2004). Teaching quality: Teacher certification and licensure. Retrieved September 5, 2004, from http://www.ecs.org/ecsmain.asp?page=/html/publications/home_publications.asp
- Ehrenberg, R., & Brewer, D. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review*, 13(1), 1-17.
- Ehrenberg, R., & Brewer, D. (1995). Did teachers' verbal ability and race matter in the 1960s? Coleman revisited. *Economics of Education Review*, 14, 1-21.
- Ehrenberg, R., Goldhaber, D., & Brewer, D. (1995). Do teachers' race, gender and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial and Labor Relations Review*, 48, 547-561.
- Ellis, S., & Worthington, L. (1994). *Research synthesis on effective teaching principles and the design of quality tools for educators* (Tech. Rep. No. 5). Eugene: University of Oregon, National Center to Improve the Tools of Educators.

- Retrieved January 12, 2005, from
<http://idea.uoregon.edu/~ncite/documents/techrep/tech05.pdf>
- Feistritzer, E. (2004). *Alternative teacher certification: A state-by-state analysis 2004*. Washington, DC: National Center for Education Information.
- Feistritzer, E., & Chester, D. (2000). *Alternative teacher certification: A state-by-state analysis 2000*. Washington, DC: National Center for Education Information.
- Felter, M. (1999). High school staff characteristics and mathematics achievement. *Education Policy Analysis Archives*, 7(9). Retrieved April 5, 2004, from <http://epaa.asu.edu/epaa/v7n9.html>
- Ferguson, R. (1998). Teachers' perceptions and expectations and the Black-White test score gap. In C. Jencks and M. Phillips (Eds.), *The Black-White test score gap* (pp. 273–317), Washington, DC: Brookings Institution Press.
- Ferguson, R., & Ladd, H. (1996). How and why money matters: An analysis of Alabama schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education*. Washington, DC: Brookings.
- Fidler, P. (2002). The relationship between teacher instructional techniques and characteristics and student achievement in reduced size classes (Publication No. 120). Los Angeles: Los Angeles Unified School District, Program Evaluation and Research Branch, Planning Assessment and Research Division.
- Flanders, N. (1960). *Teacher influence, pupil attitudes and achievement*. Minneapolis: University of Minnesota.
- Foorman, B., Fletcher, J., Francis, D., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, 90, 37–55.
- Fryer, R., & Levitt, S. (2004). Understanding the black-white test score gap in the first two years of school. *The Review of Economics and Statistics*, 86(2), 447–464.
- Gamoran, A., Porter, A., Smithson, J., & White, P. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis*, 19, 325–338.
- Germino-Hausken, E., Walston, J., & Rathbun, A. (2004). *Kindergarten teachers: Public and private school teachers of the kindergarten class of 1998–1999*. Washington, DC: National Center for Education Statistics.
- Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66, 361–396.

- Goldhaber, D. (2002). The mystery of good teaching. *Education Next*, 2(1), 50–55.
- Goldhaber, D., & Anthony, E. (2004). Can teacher quality be assessed effectively? A Working paper from the Center on Reinventing Public Education, University of Washington, Seattle. Retrieved May 17, 2004, from http://www.crpe.org/workingpapers/pdf/NBPTSquality_report.pdf
- Goldhaber, D., & Brewer, D. (1996). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources*, 32, 505–523.
- Goldhaber, D., & Brewer, D. (1997). Evaluating the effect of teacher degree level on educational performance. In W. Folwer (Ed.), *Developments in school finance*. Washington, DC: U.S. Department of Education.
- Goldhaber, D., & Brewer, D. (1998). Why should we reward degrees for teachers? *Phi Delta Kappan*, 10, 134–138.
- Goldhaber, 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., Brewer, D., and Anderson, D. (1999). A three-way error components analysis of educational productivity. *Education Economics*, 7(2), 199–208.
- Goldstein, L. (1997). Between a rock and a hard place in the primary grades: The challenge of providing developmentally appropriate early childhood education in an elementary school setting. *Early Childhood Research Quarterly*, 12(1), 3-25.
- Graham, P. (1987). Black teachers: A drastically scarce resource. *Phi Delta Kappan*, 68(8), 598–605.
- Guyton, E., Fox, M., & Sisk, K. (1991). Comparisons of teaching attitudes, teacher efficacy, and teacher performance of first-year teachers prepared by alternative and traditional teacher education programs. *Action in Teacher Education*, 13(2), 1–9.
- Hamilton, L., & Guarino, C. (2004). Conceptualization and measurement of the characteristics, practices, and philosophies of the teachers of young children. Unpublished manuscript, RAND, Child Policy Project, Santa Monica, CA.
- Hanushek, E. (1981). Throwing money at schools. *Journal of Policy Analysis and Management*, 1(10), 19–41.
- Hanushek, E. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24, 1141–1177.

- Hanushek, E. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100(1), 84–117.
- Hanushek, E. (1996). School resources and student performance. In Burtless, Gary, ed., *Does Money Matter: The Effect of School Resources on Student Achievement and Adult Success*. Washington, DC: Brookings.
- Hanushek, E., Kain, J., and Rivkin, S. (1999). *Do higher salaries buy better teachers?* Working Paper No. 7082. Cambridge, MA: National Bureau of Economic Research.
- Hardy, L. (1998). A good teacher is hard to find. *The American School Board Journal*, 9(185), 20–23.
- Hawk, P., Coble, C., & Swanson, M. (1985). Certification: It does matter. *Journal of Teacher Education*, 36, 13–15.
- Haycock, K. (2002/2003). Towards a fair distribution of teacher talent. *Educational Leadership*, 60(4), 11–15.
- Head Start Act, Public Law 105-285 (1998).
- Henke, R., Chen, X., & Goldman, G. (1999). *What happens in classrooms? Instructional practices in elementary and secondary schools, 1994–95*. Washington, DC: National Center for Education Statistics.
- Hess, F. (2001). *Tear down this wall: The case for a radical overhaul of teacher certification*. Washington, DC: Progressive Policy Institute, 21st Century Schools Project.
- Hiebert, J., Galimore, R. & Stigler, J. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher*, 31(5), 3-15.
- Holland, R. (2004, February). Competition arrives for teacher certification. *School Reform News*. Chicago, IL: Heartland Institute. Retrieved July 10, 2004, from <http://www.heartland.org/Article.cfm?artId=14357>
- Holmes Group. (1986). *Tomorrow's teachers*. East Lansing, MI: Holmes Group.
- Houston, W., Marshall, F., & McDavid, T. (1993). Problems of traditionally prepared and alternatively certified first-year teachers. *Education and Urban Society*, 26(1), 78–89.

- Huang, S., Yi, Y., & Haycock, K. (2002). *Interpret with caution: The first state Title II reports on the quality of teacher preparation*. Washington, DC: Education Trust. Retrieved July 21, 2002, from <http://www.edtrust.org/main/documents/titleII.pdf>
- Huffman, L., & Speer, P. (2000). Academic performance among at-risk children: The role of developmentally appropriate practices. *Early Childhood Research Quarterly, 15*, 167–184.
- Imig, D. (2003). *Contextual scan*. Personal statement delivered June 1, 2003, to the American Association for Colleges of Teacher Education. Retrieved July 10, 2004, from <http://www.education.armstrong.edu/context.html>.
- Ingersoll, R., (2001b). *Teacher turnover, teacher shortages, and the organization of schools*. Seattle: University of Washington, Center for the Study of Teaching and Policy.
- Ingersoll, R. (2003). Is there really a teacher shortage? A report. Seattle: University of Washington, Center for the Study of Teaching and Policy.
- Ingersoll, R., & Gruber, K. (1996). *Out-of-field teaching and educational equality*. Washington, DC: National Center for Education Statistics, Office of Educational Research and Improvement. Retrieved May 20, 2004, from <http://nces.ed.gov/pubs/96040.pdf>
- Ingersoll, R., & Smith, T. (2003). The wrong solution to the teacher shortage problem. *Educational Leadership, 60*(8), 30–33.
- Jenks, C., Smith, M., Ackland, H., Bane, M., Cohen, D., Gintis, H., Heynes, B., and Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Jepsen, C., & Rivkin, S. (2002). Class size reduction, teacher quality, and academic achievement in California public elementary schools. San Francisco: Public Policy Institute of California.
- Jordan, H., Mendro, R., & Weerasinghe, D. (1997, July). Teacher effects on longitudinal student achievement. Paper presented at the CREATE annual meeting, Indianapolis, IN.
- Kane, Parson, & Associates. (2004). A survey of principals in schools with Teach for America corps members. Unpublished report.
- Kaye, E. (Ed.). (2002). *Requirements for certification of teachers, counselors, librarians, and administrators for elementary and secondary schools, 2002–2003* (67th ed.). Chicago: University of Chicago Press.

- Klein, S., Hamilton, L., McCaffrey, D., Stecher, B., Robyn, A., & Burroughs, D. (2000). *Teaching practices and student achievement: Report of first-year findings from the "Mosaic Study" of systemic initiatives in mathematics and science* (MR-1233-EDU). Santa Monica, CA: RAND.
- Kruse, S., Seashore Louis, K., & Bryk, A. (1994). *Building professional community in schools*. Madison: University of Wisconsin, Center on Organization and Restructuring of Schools.
- Laczko-Kerr, I., & Berliner, D. (2002). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy [Electronic version]. *Education Policy Analysis Archives*, 10. Retrieved September 3, 2003, from <http://epaa.asu.edu/epaa/v10n37/>
- Levin, J., & Quinn, M. (2004). *Missed opportunities: How we keep high quality teachers out of urban classrooms*. New York: New Teacher Project, Inc.
- Lilly, M. (1992). Research on teacher licensure and state approval of teacher education programmes. *Teacher Education and Special Education*, 15, 148–160.
- Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- Lubeck, S. (1998). Is developmentally appropriate practice for everyone? *Childhood Education*, 74, 283–292.
- Lutz, F., & Hutton, J. (1989). Alternative teacher certification: Its policy implications for classroom and personnel practice. *Education Evaluation and Policy Analysis*, 11(3), 237–254.
- Mandeville, G., & Liu, Q. (1997). The effect of teacher certification and task level on mathematics achievement. *Teaching and Teacher Education*, 13, 397–407.
- Marcon, R. (1999). Differential impact of preschool models on development and early learning of inner-city children: A three-cohort study. *Developmental Psychology*, 35, 358–375.
- Marzano, R. (1992). *A different kind of classroom: Teaching with the dimensions of learning*. Alexandria, VA: Association of Supervision and Curriculum Development.
- Marzano, R., Gaddy, B., Dean, C. (2000). *What works in classroom instruction?* Alexandria, VA: Association of Supervision and Curriculum Development.
- Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, 21, 29–45.

- McBrien, J., & Brandt, R. (1997). *The language of learning: A guide to educational terms*. Alexandria, VA: Association of Supervision and Curriculum Development.
- McMillian, J. (2003). The relationship between instructional and classroom assessment practices of elementary teachers and student scores on high-stakes tests (Report No. TM034718). Virginia: ERIC Clearinghouse on Assessment and Evaluation. (ERIC Document Reproduction Service No. ED472164) Retrieved on December 17, 2004 from http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/28/04/d0.pdf
- Medley, D., & Mitzel, H. (1959). Some behavioral correlates of teacher effectiveness. *Journal of Educational Psychology*, 50, 239–246.
- Miller, J., McKenna, M., & McKenna, B. (1998). A comparison of alternatively and traditionally prepared teachers. *Journal of Teacher Education*, 49, 165–176.
- Morvant, M., Gersten, R., Gillman, J., Keating, T., & Blake, G. (1995). Attrition/retention of urban special education teachers: Multifaceted research and strategic action planning. Eugene, OR: Eugene Research Institute. (ERIC Document Reproduction Service No. ED338154)
- Munoz, M. (2000). Mathematics achievement in a reform environment: The effect of teacher, student, and parental characteristics on student testing performance (ERIC Document Reproduction Service No. ED463168).
- Murnane, R., & Phillips, B. (1981). What do effective teachers of inner-city children have in common? *Social Science Research*, 10, 83–100.
- National Association for the Education of Young Children. (2004). www.naeyc.org
- National Center for Education Statistics. (1992). *1990 science report card*. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (1996). *High school seniors' instructional experiences in science and mathematics*. Washington, DC: U.S. Government Printing Office.
- National Center for Education Statistics. (1999). *What happens in classrooms? Instructional practices in elementary and secondary schools, 1994–95* (NCES 1999-348). Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (2002). *Qualifications of the public school teacher workforce: 1987–1988 to 1999–2000*. Washington, DC: Author. Retrieved July 1, 2002, from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002603>

- National Commission on Teaching for America's Future. (1996). *What matters most: Teaching for America's future*. New York: Author.
- National Council of Mathematics Teachers. (2005). *A research companion to principles and standards for school mathematics*. Washington, DC: Author. Retrieved January 11, 2005, from http://www.nctm.org/standards/12752_exec_pssm.pdf
- National Institute for Early Education Research. (2004). www.nieer.org
- Ng, J. (2003). Teacher shortages in urban schools: The role of traditional and alternative certification routes in filling the voids. *Education and Urban Society*, 35(4), 380–398.
- Nutall, G., & Alton-Lee, A. (1993). Predicting learning from student experience of teaching: A theory of student knowledge in classrooms. *American Education Research Journal*, 30, 799–840.
- Odden, A., & Kelley, C. (2002). *Paying teachers for what they know and do: New and smarter compensation strategies to improve schools*. Thousand Oaks, CA: Corwin Press.
- Perie, M., Baker, D., & Bobbitt, S. (1997). *Time spent teaching core academic subjects in elementary schools: Comparisons across community, school, teacher, and student characteristics*. Washington, DC: U.S. Dept. of Education, Office of Educational Research and Improvement, Educational Resources Information Center, National Center for Education Statistics.
- Podgursky, M. (2001, Summer). Defrocking the national board: Will the imprimatur of “Board Certification” professionalize teaching? *Education Matters*, 1(2), 79-82. Retrieved January 18, 2005, from <http://www.educationnext.org/20012/79.pdf>
- Pressley, M., Rankin, J., & Yokoi, L. (2001). A survey of instructional practices of primary teachers nominated as effective in promoting literacy. *The Elementary School Journal*, 96(4), 363–369.
- Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *Elementary School Journal*, 83, 427–452.
- Pyburn, K. (1990). Legal challenges to licensing examinations. *Educational Measurement: Issues and Practices*, 9(4), 5–6, 14.
- Qu, Y., & Becker, B. (2003). *Does traditional teacher certification imply quality?* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

- Raymond, M., Fletcher, S. H., & Luque, J. (2001). *Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Unpublished manuscript, Stanford University, Stanford, CA.
- Recruiting New Teachers. (2000). The urban teacher challenge: Teacher demand and supply in the great city schools. Retrieved July 3, 2004, from http://www.rnt.org/channels/clearinghouse/aboutrnt/1562_puburbanteacher.htm
- Rice, J. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, DC: Progressive Policy Institute.
- Rose, L., & Gallop, A. (2003). The 35th annual Phi Delta Kappa/Gallop poll of the public's attitudes towards the public schools. *Phi Delta Kappan*, 82(1), 41-57.
- Rowan, B., Chiang, F., & Miller, R. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, 70, 256-284.
- Rowan, B., Correnti, R., & Miller, R. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects students of elementary schools. *Teachers College Record*, 104(8), 1525-1584.
- Rozycki, E. (1999). The dynamics of teacher certification: Mythologies of competition. *Educational Horizons*, 78(1), 7-10.
- Russell, T., & Munby, H. (1991). Reframing: The role of experience in developing teachers' professional knowledge. In D. A. Schön (Ed.), *The reflective turn* (pp. 164-187). New York: Teachers College Press.
- Sanders, R., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future academic achievement* Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.
- Schrag, F. (1995). *Back to basics: Fundamental education questions reexamined*. San Francisco: Jossey-Bass.
- Shen, J. (1998). Alternative certification, minority teachers, and urban education. *Education and Urban Society*, 31, 30-41.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1-22.
- Stecher, B., Hamilton, L., Ryan, G., Vi-Nhuan, L., Williams, V., Robyn, A., & Alonzo, A. (2002). *Measuring reform-oriented instructional practices in mathematics and science* (DRU-2787-EDU). Santa Monica, CA: RAND Corporation.

- Steele, C. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613–629.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States* (NCES 99-074). Washington, DC: National Center for Education Statistics.
- Stipek, D., Feiler, R., Daniels, D., & Milburn, S. (1995). Effects of different instructional approaches on young children's achievement and motivation. *Child Development*, 66, 209–223.
- Stone, J. (2002). *The value-added achievement gains of NBPTS-certified teachers in Tennessee: A brief report*. Johnson City, TN: East Tennessee State University, College of Education. Retrieved January 18, 2004, from <http://www.education-consumers.com/briefs/stoneNBPTS.shtm>
- Thompson, S. (2001). The authentic standards movement and its evil twin. *Phi Delta Kappan*, 82(5), 358–362.
- Tomlinson, C. (2001). *Differentiated instruction in the mixed-ability classroom*. Alexandria, VA: Association of Supervision and Curriculum Development.
- Tryneski, J. (1997). *Requirements for certification of teachers, counselors, librarians, administrators for elementary and secondary schools* (62nd ed.). Chicago, IL: University of Chicago Press.
- United States Department of Education (2002). *Meeting the highly qualified teachers challenge: The secretary's annual report on teacher quality*. Washington, DC: Author. Retrieved July 13, 2002, from <http://www.ed.gov/about/reports/annual/teachprep/2003title-ii-report.pdf>
- Vandevoort, L. G., Amrein-Beardsley, A. & Berliner, D. C. (2004, September 8). National board certified teachers and their students' achievement. *Education Policy Analysis Archives*, 12(46). Retrieved January 5, 2005 from <http://epaa.asu.edu/epaa/v12n46/>.
- Vartuli, S. (1999). How early childhood teacher beliefs vary across grade level. *Early Childhood Research Quarterly*, 14, 489–514.
- Von Secker, C., & Lissitz, R. (1999). Estimating the impact of instructional practices on student achievement. *Journal of Research in Science Teaching*, 36, 1110–1126.

- Walsh, K. (2001). *Teacher certification revisited: Stumbling for quality*. Baltimore: Abell Foundation. Retrieved July 3, 2002, from http://www.abell.org/pubsitems/ed_cert_1101.pdf
- Wayne, A. J. (2000). Teacher supply and demand: Surprises from primary research. *Education Policy Analysis Archives*, 8(47). Retrieved June 5, 2003, from <http://epaa.asu.edu/epaa/v8n47.html>
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10. Retrieved September 16, 2004, from <http://epaa.asu.edu/epaa/v10n12/>
- Wise, A. (2003, April 9). Commentary: What's wrong with teacher certification? *Education Week*. Retrieved May 5, 2004, from <http://www.edweek.org/ew/articles/2003/04/09/30wise.h22.html>
- Woolley, S. L, Woolley, A. W., & Hosey, M. (1999). *Impact of student teaching on student teachers' beliefs related to behaviorist and constructivist theories of learning*. Paper presented at the Association of Teacher Educators, Chicago, IL.
- Wright, S., Horn, S., & Sanders, W. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57–67.