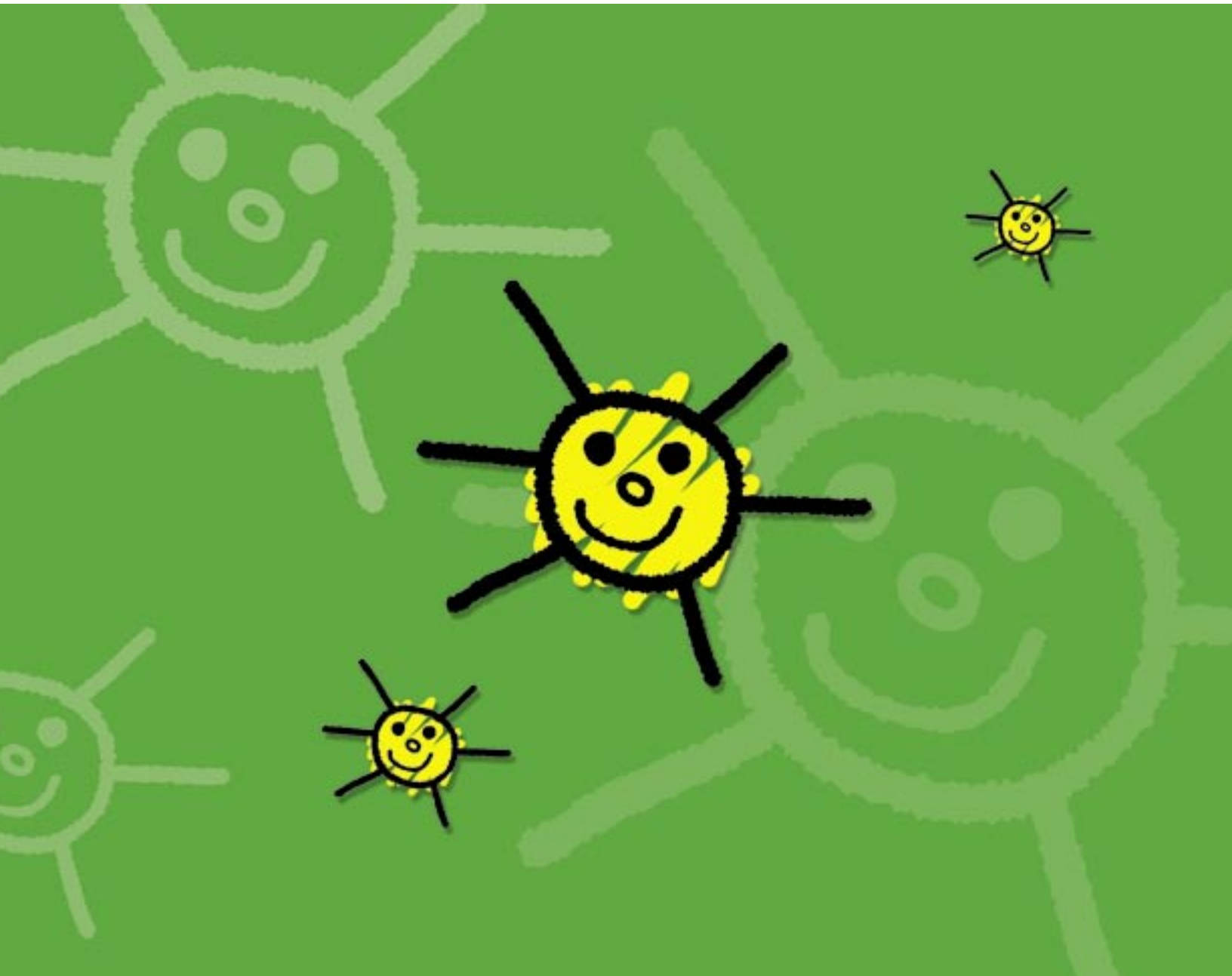
NCES
National Center for Education Statistics

Early Childhood Longitudinal Study

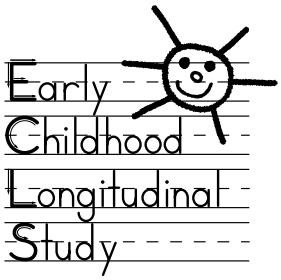
Teacher Qualifications, Instructional Practices, and Reading and Mathematics Gains of Kindergartners

Research and Development Report

U.S. Department of Education
Institute of Education Sciences
NCES 2006-031



National Center for Education Statistics



U.S. Department of Education
Institute of Education Sciences
NCES 2006-031

Teacher Qualifications, Instructional Practices, and Reading and Mathematics Gains of Kindergartners

Research and Development Report

March 2006

Cassandra M. Guarino
Laura S. Hamilton
J.R. Lockwood
RAND Corporation

Amy H. Rathbun
Education Statistics Services Institute

Elvira Germino Hausken
Project Officer
Institute of Education Sciences
National Center for
Education Statistics

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high-quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public. Unless specifically noted, all information contained herein is in the public domain.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006-5651

March 2006

The NCES World Wide Web Home Page address is http://nces.ed.gov.
The NCES World Wide Web Electronic Catalog is http://nces.ed.gov/pubsearch.

Suggested Citation
Guarino, C.M., Hamilton, L.S., Lockwood, J.R., and Rathbun, A.H. (2006). *Teacher Qualifications, Instructional Practices, and Reading and Mathematics Gains of Kindergartners* (NCES 2006-031). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

For ordering information on this report, write to
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

or call toll free 1-877-4ED-Pubs or order online at http://www.edpubs.org.

Content Contact
Elvira Germino Hausken
(202) 502-7352
elvira.hausken@ed.gov

Executive Summary

There is increasing interest among educators, policymakers, and researchers in understanding the factors that make some teachers more effective than others, particularly in light of the current focus on educational accountability at the local, state, and national levels. Thus far, only a small body of research exists, however, that links specific teacher qualifications to student achievement. The lack of research is due primarily to the scarcity of data that link student test scores to the characteristics of their teachers. Furthermore, although scholars and policymakers agree that children's early school and family experiences are pivotal, relatively little research exists on the effects of teachers on the educational outcomes of young children.

This study fills a gap in the current research base on the relationship among teacher characteristics, instructional practices, and the achievement of young children through an analysis of data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). These data were collected by the National Center for Education Statistics (NCES), within the U.S. Department of Education's Institute of Education Sciences, from a nationally representative sample of the nation's 1998–99 kindergarten class. The students were assessed in reading and mathematics in both the fall and the spring of their kindergarten year, and detailed information was gathered from their parents, teachers, and school administrators. In particular, the teachers were surveyed with regard to their background qualifications and the instructional practices they use in the classroom. As a result, ECLS-K data may provide information relevant to the relationships between teacher-reported qualifications and instructional practices and student achievement during the kindergarten year.

Data from ECLS-K were used to estimate the degree to which specific aspects of teacher training—the teaching credential and coursework in pedagogy—and teaching experience were associated with student achievement. In addition, the study identified teacher-reported instructional practices associated with student achievement gains and examined the qualifications of teachers and aspects of teacher training that were related to the use of these practices. Thus, the study addressed the following research questions:

- To what extent are kindergarten teachers' qualifications and instructional practices associated with gains in reading and mathematics of their students over the course of the kindergarten year?
- How are the instructional practices of kindergarten teachers related to their qualifications?

Using two-level hierarchical linear modeling (HLM), the first set of analyses estimated the relationship between student gains in reading and mathematics and teachers' reports of their qualifications and the specific instructional practices they used in their classrooms. The second set of analyses, also using two-level HLM, estimated the relationship between teachers' reports of their use of specific instructional practices and their qualifications. Comparisons in the text were tested for statistical significance to ensure that the differences were larger than might be

expected due to sampling. Only coefficients with a *p* value of .05 or less were identified as being statistically significant.[1]

Spending more time on subject and working within a full-day kindergarten structure were found to be associated with relatively large gains in achievement. Teacher-reported instructional practice measures designed to emphasize reading and writing skills, didactic instruction, phonics, and reading and writing activities were positively associated with reading achievement gains. Instructional emphasis on traditional practices and computation, measurement and advanced topics, advanced numbers and operations, and student-centered instruction were positively associated with mathematics achievement gains. The study provided no evidence of direct relationships between the self-reported qualifications of teachers and student achievement except for employment status. Children whose kindergarten teachers were employed part time made smaller gains in reading than those whose teachers were employed full time.

The analyses conducted in response to the second research question found evidence that certain teacher background variables—particularly the self-reported amount of coursework in methods of teaching reading and mathematics—were positively related to the teacher-reported frequency of various instructional practices that, in turn, were associated with higher student achievement. The completion of coursework in methods of teaching reading was positively associated with the use of phonics instruction, mixed-achievement grouping, student-centered instruction, and reading and writing activities. Coursework in methods of teaching mathematics was positively associated with the use of practices that emphasized numbers and geometry, advanced numbers and operations, traditional practices and computation, student-centered instruction, and mixed-achievement grouping. In addition, kindergarten teaching experience was negatively related to the use of student-centered instruction in reading and positively related to the use of mixed-achievement grouping in mathematics. Teacher certification appeared unrelated to reported instructional practices, with the exception of a positive association with an emphasis on concepts of measurement and advanced topics in mathematics.

Certain caveats should be noted. Since teachers are not randomly assigned to schools and students are not randomly assigned to teachers or schools, the relationships found in this study cannot be interpreted as causal. They are instead to be interpreted as a description of existing relationships that is reflective of the de facto distribution of teachers and children within the education system. Despite these limitations, this study utilizes a full set of control variables that help mitigate selection bias and provide valuable new information regarding the relationships between student achievement, teacher-reported instructional practices, and teacher-reported qualifications for the kindergarten population. The rich data, their nested structure, and the longitudinal nature of the assessments permit analyses that provide new information regarding existing relationships between student achievement, instructional practices, and teacher qualifications for the kindergarten population.

[1]Standard *t* test values were used to determine whether individual regression coefficients were greater than zero.

Foreword

The Research and Development (R&D) series of reports at NCES has been initiated to

1. share studies and research that are developmental in nature. The results of such studies may be revised as the work continues and additional data become available.
2. share the results of studies that are, to some extent, on the "cutting edge" of methodological developments. Emerging analytic approaches and new computer software development often permit new and sometimes controversial analyses to be done. By participating in "frontier research," we hope to contribute to the resolution of issues and improved analysis.
3. participate in discussions of emerging issues of interest to education researchers, statisticians, and the federal statistical community in general. R&D reports may document workshops and symposia sponsored by NCES that address methodological and analytical issues or may share and discuss issues regarding NCES practices, procedures, and standards.

The common theme in all three goals is that these reports present results or discussions that do not reach definitive conclusions at this point in time, either because the data are tentative, the methodology is new and developing, or the topic is one on which there are divergent views. Therefore, the techniques and inferences made from the data are tentative and subject to revision. To facilitate the process of closure on the issues, we invite comment, criticism, and alternatives to what we have done. Such responses should be directed to

Marilyn Seastrom
Chief Statistician
Statistical Standards Program
National Center for Education Statistics
1990 K Street NW
Washington, DC 20006-5651

Acknowledgments

The authors wish to thank Sandy Eyster of the Education Statistics Services Institute (ESSI) and Jerry West, formerly the Program Director of the Early Childhood and Household Studies (ECHS) Program of the National Center for Education Statistics (NCES) and currently with Mathematica Policy Research, Inc., for their helpful comments and assistance with the development of this report. We thank Lizeth Bejarano of the RAND Corporation for her clerical support and Greg Ridgeway, also with RAND, for his statistical support. We also appreciate the support provided by Elvira Germino Hausken, NCES project officer for the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), and Chris Chapman, NCES Program Director for ECLS-K. In addition, we appreciate the technical review comments provided by Marilyn Seastrom, Val Plisko, Kashka Kubzdela, and John Ralph of NCES; Lisa Bridges of the Institute of Education Sciences (IES); and Rachel Dinkes, Devin Carlson, Heather Freilich, and Stephen Mistler of ESSI. Special thanks also go to Mike Rollins of ESSI for his graphical support for the report.

Contents

List of Tables

Table **Page**

1. Introduction

There is increasing interest among educators, policymakers, and researchers in understanding the factors that may make some teachers more effective than others. The current focus on educational accountability at the local, state, and national levels underscores the importance of understanding these factors. Only a limited body of research exists that links specific qualifications of teachers to student achievement. The evidence that teachers with credentials promote student achievement more effectively than those without credentials is somewhat tenuous and mixed (Fetler 1999; Goldhaber and Brewer 2000). Some research finds small relationships between teachers' subject-matter expertise and the achievement of their students (Goldhaber and Brewer 2000; Monk 1994). One of the primary reasons for the lack of a large body of definitive research on the topic of teacher effectiveness is the paucity of data that link student test scores to the characteristics of their teachers. Furthermore, although scholars and policymakers agree that children's early school and family experiences are pivotal, relatively little research exists on the effects of teachers on the educational outcomes of young children.

This study aims to fill a gap in the existing body of research on relationships linking teacher qualifications and teaching practices to the achievement of young children through an analysis of data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K). These data were collected by the National Center for Education Statistics (NCES), within the U.S. Department of Education's Institute of Education Sciences, from a nationally representative sample of the nation's 1998–99 kindergarten class. Students were assessed in reading and mathematics in both the fall and the spring of the kindergarten year. Their parents were interviewed extensively about family circumstances and demographics. Children's teachers were surveyed regarding their qualifications and the instructional practices they use in the classroom. Also, children's school administrators were surveyed with regard to school characteristics. As a result, ECLS-K data may provide information relevant to the relationships linking self-reported teacher qualifications and instructional practices to student achievement during the kindergarten year.

In this study, data from ECLS-K are used to estimate the degree to which specific aspects of teacher training—the teaching credential and coursework in pedagogy—and teaching experience are associated with student achievement. In addition, the study identifies the teacher-reported instructional practices associated with student achievement gains and examines the types of training that are related to the use of these practices. Specifically, the study addresses the following research questions:

- To what extent are kindergarten teachers' qualifications and instructional practices associated with gains in reading and mathematics of their students over the course of the kindergarten year?
- How are the instructional practices of kindergarten teachers related to their qualifications?

The relationships revealed in this study cannot be interpreted as causal due to the lack of random assignment of children to teachers and schools. However, this study addresses unanswered

questions regarding the relationships linking the qualifications and instructional practices of kindergarten teachers to achievement outcomes. This report is organized as follows. Section 2 discusses the existing research on teacher effectiveness and instructional practices. Section 3 describes the data, and section 4 presents the methods used in this study. Section 5 presents the findings. Section 6 discusses the findings as well as possibilities for extending this research.

In addition, the report contains five appendixes. Appendix A provides a description of the ECLS-K direct cognitive assessment measures. Appendix B provides a description of the methodology used to create the instructional practice measures used in this study. Appendix C provides a description of the methodology used in the regression analyses. Appendix D presents sample statistics for variables used in this study. Appendix E presents the standard errors associated with the coefficients in the regression analyses included in the main text of the report.

2. Background

Recent studies have suggested that teachers exert a significant influence on student learning. In particular, studies with elementary school students using value-added modeling approaches report that the overall effects of teachers may be large (Rowan, Correnti, and Miller 2002; Wright, Horn, and Sanders 1997). If the findings of large teacher effects are accurate, they suggest that efforts to improve education for all students must seek to identify the characteristics that distinguish effective from ineffective teachers. However, the research literature devoted to detecting the effects of specific teacher qualifications is inconsistent and sparse. Furthermore, few nationally representative studies have focused on relationships between teacher qualifications, instructional practices, and children's achievement in kindergarten. This section provides a brief discussion of existing research on the effects of teacher qualifications on student achievement. This discussion is followed by a summary of the literature that has examined relationships between specific instructional practices and student achievement.

Characteristics of Effective Teachers

Studies that have examined available indicators of teacher preparation or quality—such as academic ability, certification status, subject-matter expertise, and experience—offer mixed findings, suggesting that there is not yet a consensus as to what characteristics influence achievement. Prior research focusing on the impact of credentials and preservice training (e.g., studies to obtain a teaching credential) on the quality of instruction has found inconsistent results regarding the impact of the teaching credential and small positive effects regarding subject-matter preparation. Hawk, Coble, and Swanson (1985) found that for secondary school students, mathematics achievement was positively associated with having a teacher who was fully certified in mathematics. Fetler (1999) found a negative correlation at the school level between mathematics scores and the percentage of teachers with emergency credentials. In an analysis using individual student data, however, Goldhaber and Brewer (2000) found that the performance of high school students on standardized mathematics and science tests did not differ according to whether their teachers held standard or emergency credentials. They found, however, that students of teachers who were uncertified or who held a private school certification in mathematics had somewhat lower achievement levels than students of teachers with a standard, probationary, or emergency certification. Rowan, Correnti, and Miller (2002) found no effect of certification status on achievement growth at the elementary level. One caveat to be applied to studies dealing with teaching credentials is that the credential is not a standard measure of preparation, because requirements vary widely from state to state. Thus, the type of credential a teacher holds is a somewhat imperfect measure of preservice training.

With regard to subject-matter preparation, Monk (1994), in a study of secondary high school students that controlled for prior test scores, found a positive effect of teacher coursework in both college mathematics and mathematics pedagogy on mathematics test scores. Goldhaber and Brewer (2000) found a weak positive association between high school achievement gains in mathematics and teachers with a master's degree in mathematics versus those without a master's degree or a mathematics-related bachelor's degree. A recent study focusing on elementary school

students found no significant evidence of an impact of advanced degrees on achievement gains (Rowan, Correnti, and Miller 2002).

With regard to teaching experience, Fetler (1999) found a positive relationship between the average number of years of teaching experience of mathematics teachers and high school mathematics scores at the school level. Rowan, Correnti, and Miller (2002) found a positive relationship between teaching experience and growth in mathematics achievement for students as they progressed from third to sixth grade. Murnane and Phillips (1981), in a study of the achievement of Black inner-city elementary school students that controlled for prior-year test scores, found a nonlinear relationship between achievement and teaching experience. For students in grades three, four, and six, having teachers with 15 or more years of experience was positively associated with gains. For fourth- and sixth-grade students, having teachers with 7 or fewer years of experience was also positively associated with gains. For students in the third grade, having teachers with 8 to 14 years of experience was negatively associated with gains.

ECLS-K data provide information on several aspects of teacher training and qualifications, including teachers' reports of their certification status, years of kindergarten teaching experience, and advanced degree status. In addition, the data provide valuable information not normally collected by other surveys, such as the amount of training teachers have received in methods of teaching reading and mathematics and the categories of instructional practices they use. Thus, this study was able to follow, in many cases, and extend the approaches used in prior studies with older children. In addition, ECLS-K data provide an opportunity to explore the relationship between teachers' self-reported qualifications and instructional practices, as well as the links between each of these categories of factors and student achievement.

Instructional Practices and Student Achievement

To the extent that any of the characteristics explored in the previous section influence student learning, it is likely that they exert their effect in large part via the instructional practices that teachers use in the classroom. Early childhood reading researchers recommend that attention be given in every primary-grade classroom to a wide array of early reading skills including the alphabetic principle, reading sight words, reading words by mapping speech sounds to parts of words, achieving fluency, and comprehension (Snow, Burns, and Griffin 1998). In kindergarten, Ball and Blachman (1991) found that children who received instruction in phonemic segmentation and letter-sound combinations had higher reading and spelling scores than children who did not receive such instruction. While there is general agreement on the skills that should be taught, approaches on how best to teach reading are often debated. Some research with primary-grade children has shown that no single approach to teaching reading is superior to the rest, but that balanced or integrated early literacy instruction (i.e., instruction that includes an emphasis on both phonetics and meaning) is more effective in learning to read (Adams 1990; Chall 1992; Snow, Burns, and Griffin 1998). Certain teaching strategies and methods have also been found to be associated with greater gains in early mathematics learning. A meta-analysis of studies including kindergarten research found that the use of concrete materials or manipulatives, compared with more abstract instruction, was related to improved achievement and attitudes toward mathematics (Sowell 1989), especially during initial exposure to a skill (Baroody 1989). Also, problem-solving activities and opportunities to practice what is learned have been shown to

contribute to learning mathematics concepts and skills in the first years of school (National Research Council 2001).

The measurement of instructional practices, however, is less straightforward. Most studies that examine classroom practices and the links between these practices and achievement have relied exclusively, or primarily, on teacher self-reports of the frequency with which teachers use specific practices such as cooperative learning groups, inquiry-based activities, manipulatives, and open-ended assessment techniques.

Several of the studies that have used this method of measuring practices have examined links between student achievement and teacher-reported use of practices that emphasize problem solving and inquiry. These practices are often called "reform based" or "standards based" and include such activities as cooperative learning groups, student-led discussions, and open-ended assessment techniques that are intended to promote the development of complex cognitive skills and processes (Cohen and Hill 2000; Hamilton et al. 2003). Cohen and Hill's (2000) study of teacher-reported use of reform-based practices found that frequency of use was positively related to mathematics test scores among fourth-graders. D'Agostino (2000) found that in first and second grades, teacher-directed basic skill orientation was associated with greater reading and mathematics achievement, while in fourth grade, a combination of basic skill instruction and student-centered advanced skill activities was associated with greater mathematics achievement. A synthesis of data from 11 National Science Foundation-funded Systemic Reform Initiatives found a mixture of null and small positive results on both multiple-choice and open-response assessments for upper elementary and middle school teachers and students (Hamilton et al. 2003). Several studies have attempted to examine whether the relationship between achievement and reform-based practices is different from the relationship between achievement and more traditional forms of instruction, such as the use of textbooks, lectures, and multiple-choice tests (Cohen and Hill 2000; Hamilton et al. 2003). Most of the differences in the relationships between student achievement and each of these types of practice have been found to be small.

The small magnitudes of observed relationships may be attributable in part to measurement problems. Limited evidence exists regarding the extent to which survey-based measures of practices accurately capture what teachers do in the classroom. Stipek and Byler (2004) recently compared observational data of teacher instructional practices in kindergarten and first-grade classrooms with teachers' self-reported practices, using instructional practice items that were similar in content to the ECLS-K item set, and found that teachers' reports of their instructional practices were significantly correlated with observations of their practices. Mayer (1999) reports results of a study that compared information obtained from survey responses to information collected from classroom observations of middle school and high school algebra teachers. Results from his study indicated that a composite of teachers' emphasis on reform-based mathematics teaching practices based on survey data was strongly correlated ($r = .85$) with a parallel composite based on classroom observations of the teachers' mathematics instruction. He found that the survey data could be used to distinguish teachers who engaged in low levels of reform-based practices from those who relied heavily on such practices, but his study also suggested that survey results masked potentially important differences among teachers in how they implemented various practices. However, Mayer (1998) noted that when composites of individual teacher practice variables from one-time teacher surveys were grouped together to describe teachers' pedagogical style, the reliability of the resulting composites was quite high. As

a result of these measurement issues, the observed weak effects between instructional practices and achievement may result from superficial implementation by teachers (which is not captured by frequency-based survey responses) rather than from the practices' ineffectiveness. In addition, the kinds of instructional practices measured by most surveys constitute only a small part of the classroom instructional environment. Other aspects, such as the specific curriculum materials used and the pacing of instruction, are often not measured.

Despite the limitations of paper-and-pencil frequency-based measures of practice, they currently provide one of the best sources of information about the distributions of instructional practices used by teachers across the nation. Surveys provide a cost-effective and reliable method for gathering data on practices across different state and district contexts.

Most of the existing research on practices has focused on teachers at the higher elementary and secondary grades. Because of the growing recognition of the importance of developing literacy and quantitative skills from the earliest grades, there is a need to extend the work on instructional practices to include kindergarten and other early elementary grade levels. ECLS-K data on the instructional approaches used by kindergarten teachers permit this type of exploration. Xue and Meisels (2004) used these data to examine several classroom-level predictors of reading achievement and found that teachers' reported use of a phonics-based approach, in particular, was related to achievement in reading (effect size of 0.06). The research described in this report builds on this work and extends it to mathematics.

3. Data

ECLS-K data provide detailed information on a nationally representative sample of 22,000 children enrolled in approximately 1,000 kindergarten programs, both public and private, in the United States during the 1998–99 school year. The two waves of data utilized in this study were collected in the fall and spring of the 1998–99 kindergarten year. This study uses five categories of data: achievement assessments, student characteristics, teacher qualifications, instructional practices, and school characteristics.

Achievement Assessments

Assessments that included cognitive, psychomotor, and physical components were conducted with the sampled kindergarten children through one-on-one tests administered by trained individuals both at the beginning and at the end of the school year.[1] Having both fall and spring measures is one of the features that distinguishes ECLS-K data from other longitudinal study data and makes them uniquely suited for an investigation of teacher effects over the course of the kindergarten year.

The full achievement assessment used a computer-assisted personal interview (CAPI) and took approximately 50–70 minutes to complete. It included tests of reading, mathematics, and general knowledge. The test was untimed and required children to respond verbally or through pointing; no writing was required. Each test was conducted using a two-stage design, known as adaptive testing. The first stage consisted of a routing section that was administered to all students, and the second stage consisted of one of several alternative forms, the choice of which depended on (i.e., was adapted to) the child's performance on the first stage. Although the assessments included a general knowledge test, only the assessments in reading and mathematics were used in this study. The reading and mathematics assessments had low-, middle-, and high-difficulty second-stage adaptive options. The purpose of this adaptive design was to maximize accuracy of measurement and minimize administration time.[2]

The content of the reading and mathematics assessments was selected to represent cognitive skills that are typically taught in kindergarten and that are thought to be important for the development of later proficiency (Rock and Pollack 2002). The reading test consisted of items measuring basic skills (e.g., letter recognition), vocabulary, and comprehension. The mathematics test covered five content strands: number sense, properties, and operations; measurement; geometry and spatial sense; data analysis, statistics, and probability; and patterns, functions, and algebra. Items tapping all five content strands are included in the kindergarten assessment, although the first strand (number sense, properties, and operations) is most heavily

[1]More detail about the ECLS-K achievement assessments is provided in appendix A of this report.
[2]See the *ECLS-K Base-Year Public-Use Data Files and Electronic Codebook: User's Manual* (Tourangeau et al. 2001) or the *Early Childhood Longitudinal Study—Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for Kindergarten through First Grade* (Rock and Pollack 2002) for a more complete description of the assessment procedures.

represented. Examples of the types of skills involved in this strand at the kindergarten level include counting sets of objects and solving simple addition problems using pictures. Examples of skills included in the measurement strand include comparing the length of two objects and identifying the value of different coins. Examples of skills in the geometry and spatial sense strand include naming simple shapes and identifying lines of symmetry. Skills in the data analysis, statistics, and probability strand include reading information from a bar graph. Examples of skills in the patterns, functions, and algebra strand in the kindergarten assessment include recognizing and extending patterns and sequences.

Efforts were made to include children who spoke a language other than English in the assessments. Prior to administering the assessments, a language-screening test—the Oral Language Development Scale (OLDS)—was administered to those children identified from their school records (or by their teachers if no school records were available) as coming from homes in which the primary language spoken was not English. Children whose performance exceeded an established cut score on the OLDS received the full English direct assessment in both reading and mathematics. Those who did not pass the OLDS were not given the reading assessment. Students who did not pass the OLDS but who spoke Spanish were given a translated form of the mathematics assessment. Thus, the size of the sample of children who took the mathematics assessment was larger than that of children who took the reading assessment. Various methods were used to confirm that the psychometric properties of the Spanish mathematics assessment were comparable to those of the English version (Rock and Pollack 2002).

This report uses children's fall and spring kindergarten overall achievement scores in reading and mathematics, which are Item Response Theory (IRT) scale scores. Children's IRT scale score values are a reflection of the number of items they would have answered correctly on the assessment if they had been administered all items in the battery. The IRT scale scores may be used to calculate longitudinal measures of gains because the scores from each round of data collection are set on a common scale. From the IRT scores, gain scores were created (in the form of a simple difference) to represent the change in relative achievement of the child over the course of the kindergarten year. All student-level models controlled for the time lapse between fall and spring test administrations, since not all children were assessed on the same day.

Student Characteristics

Parent interviews were the source of the demographic information utilized in this study. Approximately 18,950 parent interviews were completed. Parents reported the race, ethnicity, age, and sex of each child and indicated whether the child was repeating kindergarten in the 1998–99 school year. Parents also provided information about their household income, the mother's (or female guardian's) educational attainment and occupation, and the father's (or male guardian's) educational attainment and occupation. NCES used the information from these variables to create a socioeconomic status (SES) composite variable, which is standardized with a mean of 0 and a standard deviation of 1. Details on the construction of this variable are available in the user's manual for the ECLS-K base-year public-use data file (Tourangeau et al. 2001). In addition, parents identified the primary language spoken in the home.

Teacher Qualifications

Information on 3,305 kindergarten teachers was gathered in a set of self-administered paper-and-pencil questionnaires that included questions on their qualifications and instructional practices. Measures of qualifications used in this study were years of kindergarten teaching experience, employment status (i.e., part time or full time), level of certification, educational attainment (i.e., receipt of a master's degree), and the number of courses completed in methods of teaching reading and mathematics.

With regard to certification, ECLS-K data include information about the type of teaching certificate teachers reported they possessed (none; temporary, probational, provisional, or emergency; alternative; regular but less than the highest available; and the highest available—permanent or long term). The analyses in this study distinguish between teachers who have a regular or the highest certification available and those who do not. These categories were chosen on the basis of the distribution of certifications among teachers given that the vast majority of kindergarten teachers in the sample were certified and few teachers fell into the alternatively credentialed or noncertified categories (only 2 percent of teachers held alternative credentials and another 2 percent were uncertified).

One of the difficulties inherent in research on national samples is the state-specific nature of many education programs and policies. It is important to note here that requirements for teacher certification often differ by state. Paths for attaining each type of certification vary across states in a number of respects, including the standards for admission into certification programs and the amount of in-classroom experience required. Therefore, ambiguities associated with certification status must be kept in mind when interpreting results from this study.

ECLS-K data also include teacher-reported information on the area in which teachers are certified (elementary, early childhood, or other). The nature of early childhood and elementary certification is particularly variable across states, and it is common for teachers to possess both (e.g., by supplementing a standard elementary certification with a special early childhood certification). Because of this variation and the lack of clear definitions of early childhood and elementary certification, this information is not used in this study. As a result, readers are cautioned about the conclusions that can be drawn from results of the relationships between certification status and achievement gains.

Instructional Practices

The spring teacher questionnaire included a set of items that addressed instructional activities and skill emphasis in reading, mathematics, science, and social studies. The first two subjects are the focus of this study. The items address a wide range of instructional practices that may occur in classrooms in the early grades, not just practices considered to be developmentally appropriate for kindergarten.[3] Both the mathematics and reading practice items were selected to be aligned

[3]NCES and the authors do not speculate on whether certain types of practices are developmentally appropriate or inappropriate for kindergarten children. Instead, this report describes how frequently teachers use different types of instructional practices in their classrooms.

with the skills tested by ECLS-K achievement assessments. Many of the ECLS-K instructional practice items are similar in content to other research that uses teacher questionnaires to collect data on instructional practices in kindergarten classrooms (Stipek and Byler 2004).

The items included in the analysis are of two types. One set of items asked teachers, "How often do children in this class do each of the following READING and LANGUAGE ARTS activities?" Teachers were asked to rate the frequency of each activity, using a scale that ranged from 1 ("never") to 6 ("daily"). Intermediate scale points were 2 ("once a month or less"), 3 ("two or three times a month"), 4 ("once or twice a week"), and 5 ("three or four times a week"). The other set asked, "For this school year as a whole, please indicate how each of the following READING and LANGUAGE ARTS skills is taught in your class(es)." For each specified skill, teachers could indicate that the skill was not taught, by selecting either "taught at a higher grade level" or "children should already know," or could indicate the frequency with which the skill was taught, using a 5-point scale that ranged from "once a month or less" to "daily." Intermediate scale points were the same as for the activities items. Similar sets of questions were asked about mathematics instruction.

As discussed above, for the items that addressed skills taught in reading and mathematics, teachers could report that a particular skill was not taught by choosing either "taught at a higher grade level" or "children should already know." For the analyses discussed in this report, which focus on the frequency with which teachers emphasize particular skills or topics, these two options were combined into a single "not taught" category. The numerical values for all of the options were then recoded so that the scale for the "skills" items was similar to that for the "activities" items (i.e., the category representing the greatest frequency received a score of 6, and the "not taught" category received a score of 1). This procedure enabled the two sets of items to be combined into a single factor analysis. Although the difference between the two "not taught" options may be of substantive interest for some purposes, because the focus of this research was on the amount of exposure students received, combining these categories was appropriate. Therefore, item means and factor analyses are based on sets of items with response options ranging from 1 to 6. The development of instructional practice scales from these items relies on prior work conducted for NCES (Hamilton and Guarino 2005). That study used factor analyses to create several scales similar to those used in Cohen and Hill (2000) and Hamilton et al. (2003).

As a result of the factor analyses, seven scales were derived for reading and six scales were derived for mathematics. The seven reading scales were *phonics; reading and writing skills; reading and writing activities; didactic instruction; comprehension; student-centered instruction;* and *mixed-achievement grouping.* The *phonics* scale includes items addressing teachers' emphasis on activities such as learning and practicing letters and the importance teachers place on skills such as matching letters to sounds and recognizing letters. It is similar to the phonics scale used by Xue and Meisels (2004). The *reading and writing skills* scale measures teachers' reported emphasis on promoting students' ability to read fluently and to write using standard conventions (e.g., capitalization, punctuation, and conventional spelling). The *reading and writing activities* scale focuses on specific activities in which students engage to promote reading and writing ability, including reading aloud and writing in journals. *Didactic instruction* includes three items that appear to measure more traditional language arts instructional activities that students probably do as seatwork—reading from basal texts, working in workbooks or on

worksheets, and writing words from dictation to improve spelling. *Comprehension* covers skills involved in understanding what is read (e.g., identifying main ideas, making predictions, and using context cues). *Student-centered instruction* appears to address activities in which students are responsible for producing something and sharing it with the class (e.g., publishing their own writing, performing plays, and doing an activity or project related to a book or story). The final scale, *mixed-achievement grouping*, consists of two items that assess the frequency of mixed-achievement grouping and peer tutoring. Internal consistency reliability for the seven reading practices scales ranged from .69 to .83, with two exceptions. The reliability of the three-item *didactic instruction* scale was .50, and the reliability of the two-item *mixed-achievement grouping* scale was .48.

The six mathematics instructional practice scales were *numbers and geometry; advanced numbers and operations; traditional practices and computation; student-centered mathematics instruction; mixed-achievement grouping;* and *measurement and advanced topics.* Some of the scales focused on particular content, whereas others related more to the structure of classroom instruction. The first scale, *numbers and geometry*, consists of 11 items, 3 of which address classroom activities designed to promote understanding of numbers, operations, and geometry. The other eight items in this scale address instructional objectives related to numbers and geometry. The second scale, *advanced numbers and operations*, contains five items from the "skills" item set. In contrast to the first scale, these items emphasize skills related to somewhat advanced numeracy concepts (e.g., reading two- and three-digit numbers). *Traditional practices and computation* appears to identify teachers who take a fairly didactic, teacher-directed approach to instruction and who emphasize computational facility. It includes the use of worksheets, textbooks, and problems on the chalkboard and emphasizes computational skills such as adding and subtracting single-digit numbers. The fourth scale, *student-centered mathematics instruction*, is similar to the corresponding scale identified for reading, and consists of six items that emphasize having students take an active role in their own learning (e.g., explaining how a mathematics problem is solved, working on problems that reflect real-life situations, and using creative movement or creative drama to understand mathematics concepts). The fifth scale, *mixed-achievement grouping*, consists of two items: use of mixed-achievement groups and peer tutoring. Finally, *measurement and advanced topics* includes one item from the "activities" set—work with rulers, measuring cups, spoons, or other measuring instruments—and eight "skills" items emphasizing topics other than the basic numbers and geometry topics addressed by the first and second scales. These include fractions, data collection, probability, and estimation, along with measurement-specific skills such as "using measuring instruments accurately." Internal consistency reliability for the six mathematics practices scales ranged from .73 to .83, with one exception. The reliability of the two-item *mixed-achievement grouping* scale was .56.

After identifying satisfactory factor solutions in reading and mathematics, scales for each teacher were created to represent the resulting set of factors. A number of methods may be used to combine information from a set of items into a single scale. The most common method for combining items that use Likert scales is to add the scores on the individual items, a process that is typically referred to as summative scoring (or sum scoring). When the resulting scales are based on a factor analysis, a common approach is to identify the factor on which each item has its highest loading and include that item in the calculation of a score for that factor. For this report, because some scales are constructed from more items than others, scores on individual

items composing the scale were averaged, rather than totaled, for each teacher. For teachers who answered some, but not all, items on a given scale, the average score was calculated based on those items that they did answer. For teachers who did not respond to any of the items used to create a particular scale, the imputation was carried out at the scale level. For these teachers, an imputation procedure that tried to reconstruct the item-level data prior to scale creation would be difficult to implement because of the potentially large number of individual items that would have to be imputed. Imputation on the aggregate scale reduces the missing information down to a missing scalar rather than an entirely missing vector.

A summary of the methodology used in the development of the instructional practice scales and a list of items included in the scales is presented in appendix B. Descriptive statistics for the scales are presented in table D-2 of appendix D. It is important to note that these instructional practice scales provide information about teacher-reported frequency and do not address variations in intensity of instruction, total time spent in a given day, or overall quality of instruction. Therefore, there are important differences among teachers and their instructional practices that these scales cannot capture and that should be the focus of future research.

School Characteristics

Data were collected from school administrators through a self-administered questionnaire that provided information on the fiscal, organizational, and instructional environment of the school. For the purposes of this study, information was included on the size of the school, its percentage of minority students, its percentage of free-lunch-eligible students,[4] and whether it is public or private. The study also utilized geographic information relating to region and urbanicity.

[4]The variable for percentage of free-lunch-eligible students in the school has a relatively high number of cases with missing data (26 percent at the school level). The variable was imputed for the missing cases, as were all variables in the analyses.

4. Methods

The research presented in this study was conducted in two stages. The first stage consisted of using hierarchical linear modeling (HLM) to estimate the relationship between student gains in reading and mathematics achievement over the course of the kindergarten year and teachers' reports of their qualifications and the specific instructional practices they used in their classrooms. The second stage involved using HLM to estimate the relationship between teachers' reports of their use of particular instructional practices and their qualifications.

To allow for the interpretation of HLM regression coefficients as effect sizes, certain transformations of the data were carried out. To maximize the number of cases included in the analyses and to reduce bias and increase the efficiency of parameter estimates, missing data items were imputed. The modeling of the two stages of the analysis and the transformation and imputation procedures are discussed below.

Statistical Models

The modeling of the relationship between the achievement gains of kindergarten students and the characteristics and instructional practices of their teachers is aided by four characteristics of the dataset. First, ECLS-K assesses children at the time when they enter the formal K–12 education system and follows them through their early school experiences, thus allowing the researcher to measure changes over time. The achievement regressions use a gain score consisting of the difference in scores on the fall and spring assessments (as described in section 3) as the dependent variable. This approach allows for an analysis of change during the time the student was with the teacher, controlling for student background characteristics. The ability to use an achievement gain score rather than a static one-time score improves the modeling considerably by controlling for initial differences in achievement. Gain scores "difference out" unobserved child-level characteristics that have time-invariant effects on achievement. Unobserved child-level characteristics might be, for example, innate intelligence, attention span, or "enthusiasm for learning." Second, the contextual variables provided in the dataset provide a fairly comprehensive set of controls, including age, socioeconomic status, and race/ethnicity. Third, the dataset includes a set of teacher-reported qualifications and practice variables that allow researchers to explore the relative contributions of a larger number of factors that might influence student outcomes. Finally, the data are collected in a nested structure, with children nested within teachers and teachers nested within schools. The nested structure of the data allows researchers to relate student outcomes to observable features of the higher level units to which they are linked, but it is important to note that this structure can also induce spurious positive correlations among outcomes from students sharing a teacher or school. These correlations, if not properly accounted for, can reduce the efficiency of the estimates and also introduce bias into the estimated standard errors for these estimates. Thus, this study uses HLM to account for these correlations and improve the validity of inferences. In addition to accounting properly for clustering at the school and teacher levels, HLM also allows for the modeling of effects at various levels.

In the first stage of the analysis, a two-level hierarchical linear model is used to estimate relationships between students' achievement gains in reading and mathematics and teachers' reports of their qualifications and use of specific instructional practices, taking into account the clustering of students within particular teachers and schools. The dependent variables in the achievement analyses are gain scores that represent the differences between the IRT fall and spring achievement scale scores for the reading and mathematics assessments. These analyses therefore estimate the association of various factors with the amount of learning children demonstrated during the kindergarten year.[5] Mean achievement gain scores were estimated as conditional upon a set of child-, teacher-, and school-level characteristics. These analyses permit inferences regarding the association of teaching practices and teacher qualifications with achievement gains. These models include random teacher effects, which capture both the teacher- and school-level variability. Since the number of teachers per school in the data is relatively small, computational constraints permit estimating only two levels (the child and teacher levels), rather than three levels (including the school level). Accounting for clustering at the teacher level facilitates valid inferences about the variables contained in the model, which is the primary goal in the conditional means models. A number of child-level control variables and teacher and school characteristics are also included in the model, mitigating the problem of selection bias to some degree. Selection bias can occur if the choices that parents make—such as the choice of a public or private school, full- or half-day kindergarten, or a particular neighborhood based on accessibility to a certain school—result from family characteristics that are not observed but are correlated with the observed factors influencing achievement. Again, it is important to bear in mind that none of these models can produce estimates that can be interpreted as truly causal effects due to the lack of random assignment at any level. A more complete description of the hierarchical linear models can be found in appendix C.

The teacher is the unit of analysis in the second part of the analysis, which uses a similar two-level hierarchical linear model to estimate the relationship of teacher scores on the specific instructional practice scale with their self-reported qualifications, taking into account the clustering of teachers within schools. In these models, the random intercept picks up school-level effects. The type of two-level model described above applies, with the substitution of teachers for children and schools for teachers. The characteristics of teachers used as covariates in the analysis are years of experience in teaching kindergarten; the number of courses completed in methods of teaching reading or mathematics; and whether they hold full teaching certification, possess a master's degree, and work full or part time.[6] All HLM analyses applied survey weights to the level 1 variables.[7]

---

[5]An alternative approach would be to estimate the conditional expectation of spring IRT scale scores and include fall scores as independent (predictor) variables. The interpretation of the results of this type of specification differs slightly from the interpretation of the gain-score model in that it represents the status of children's learning adjusted by their fall status rather than by gains in learning. See appendix C for a discussion of studies supporting the use of gain scores.

[6]The part-time employment indicator variable was created directly from a variable called YKFULL provided in the ECLS-K restricted-use salary and benefits data file (NCES 2001-014).

[7]The child-level analyses used the weight BYCW0, and the teacher-level analyses used the weight B2TW0.

Data Transformations and Sample Sizes

Transformations of the data were made to facilitate the interpretation of coefficients. All continuous variables, both dependent and predictor (independent), that were not percentages were standardized (i.e., were scaled to have a mean of 0 and a variance equal to 1). The regression coefficients for continuous predictors may be interpreted as effect sizes and represent the number of standard deviations of the dependent variable associated with one standard deviation change in the predictor, after controlling for the other variables in the regression. All percentages (i.e., the percentage of minority students in the school and the percentage of students eligible for free lunch in the school) were on the scale of 0 to 1, so the regression coefficients represent standardized changes in the dependent variable associated with moving from 0 to 100 percent on the predictor. Dichotomous variables were not standardized and were coded with the usual 0/1 values, so that coefficients for such variables can be interpreted as the number of standard deviations of the dependent variable associated with moving from 0 to 1 on the dichotomous predictor.

The initial ECLS-K sample contained approximately 22,000 kindergarten children. The weighted response rate for schools in the base year (i.e., kindergarten) was 74 percent, with 944 of the 1,277 originally sampled schools participating. The base-year survey completion rates for children, parents, and teachers were 92, 89, and 94 percent, respectively.[8] This study retains only the subset of children that have the data requisite for the analyses: 16,308 students linked to teachers and schools who had observed achievement gain scores for at least one of the two subject assessments and who did not change teachers or schools during the course of the kindergarten year. Of these students, 15,494 had observed reading gain scores and 16,284 had observed mathematics gain scores. The difference in the size of the two samples stems from the fact that children from non-English-speaking homes who did not meet the cutoff score on the OLDS screening test but spoke Spanish were given a mathematics test in Spanish but no reading test. Rock and Pollack (2002) present results of analyses that demonstrate the equivalence of the Spanish and English versions of the mathematics test. An earlier report revealed that estimates of the average reading achievement of kindergartners are not significantly affected by the exclusion of the students who could not be tested in English (Denton and West 2002).

The teacher-level dataset contains 3,305 teachers. In the teacher-level regressions, the numbers of teachers with nonmissing values on the various reading and mathematics instructional practice scales were as follows: *phonics* (3,084), *reading and writing skills* (3,068), *reading and writing activities* (3,084), *didactic reading instruction* (3,078), *reading comprehension* (3,061), *student-centered reading instruction* (3,082), *mixed-achievement reading grouping* (3,066), *numbers and geometry* (3,073), *advanced numbers and operations* (3,057), *traditional practices and computation* (3,075), *student-centered mathematics instruction* (3,062), *mixed-achievement mathematics grouping* (3,050), and *measurement and advanced topics* (3,073).

---

[8] A nonresponse bias analysis was conducted to determine if substantial bias had been introduced due to school nonresponse. Details on the nonresponse bias analysis are included in the NCES report *Kindergarten Teachers: Public and Private School Teachers of the Kindergarten Class of 1998–99* (Germino Hausken, Walston, and Rathbun 2004). Overall findings from the analysis indicated no strong biases due to school nonresponse.

In order to minimize the impact of missing values in the covariates, missing data were imputed for the independent variables. The dependent variables were not imputed. The imputation procedures are described in appendix C. The set of sample means and other statistics for the child- and teacher-level datasets can be found in appendix D.

5. Findings

This section provides a description of the hierarchical linear modeling (HLM) regression findings for reading and mathematics. The results of the student-level achievement regressions provide evidence regarding the relationships between the achievement gains made by kindergarteners and teachers' reports of their qualifications and the instructional practices they used in their classrooms. The results of the teacher-level instructional practice scale regressions explore the relationship between instructional practices and teacher characteristics. In interpreting the regression results, alpha = .05 is the criterion for statistical significance. Therefore, only coefficients with a *p* value of .05 or less are interpreted.[9] These are marked in the tables with an asterisk.

Reading Achievement Gain Regressions

The regressions of fall-to-spring gains in reading achievement revealed several significant relationships with teacher-, school-, and child-level characteristics. Teacher-level variables present in every regression include an indicator for whether the kindergarten class was a full-day class and a measure of the time the teacher reported spending each day on reading or mathematics (tables 1 and 2). School-level variables present in every regression include school locale, control, enrollment size, percentage of minority students and percentage of free-lunch-eligible students, and region. Child-level variables present in every regression include gender, race/ethnicity, age at fall kindergarten assessment, days elapsed between assessments, kindergarten repeater status, and primary home language.

Separate regression models were run for each of the individual instructional practice scales described in section 3 of this report to gain a sense of the relationship of each of these types of practices to achievement gains (table 1). An alternative specification that included all of the teacher practice variables simultaneously is presented in the last column of table 1. A final specification excluded the instructional practice scales and instead included several variables related to the self-reported qualifications of teachers: years of total kindergarten teaching experience (with a squared term to detect nonlinear effects), certification level (i.e., full vs. less than full), employment status (i.e., full vs. part time),[10] degree status (i.e., receipt of a master's degree), and the number of courses taken in methods of teaching reading or mathematics (0 to 6 or more courses) (table 2).

Teacher-Level Variables: Kindergartners' reading gains were related to the amount of time spent in school, on subject, and on teachers' reports of their use of various instructional approaches and were unrelated to the types of self-reported teacher qualifications examined in this study. Students in full-day kindergarten programs and students whose teachers reported spending 90 minutes or more per day teaching reading exhibited significantly higher gains than those in part-

[9]Standard *t* test values were used to determine whether individual regression coefficients were greater than zero.
[10]Part-time employment status does not necessarily provide the same information as the indicator for full-day kindergarten. Teachers might work full time and teach two half-day kindergarten classes, for example.

day kindergartens and those whose teachers spent less time each day on reading instruction, with effect sizes ranging from 0.07 to 0.10 for reading instruction time and 0.16 to 0.19 for full-day kindergarten.[11] Students whose teachers reported a greater emphasis on *reading and writing skills*, a *didactic* approach, *phonics*, and *reading and writing activities* exhibited greater achievement gains than those whose teachers spent less time on such practices, with effect sizes of 0.10, 0.09, 0.07, and 0.05, respectively. Teacher use of *comprehension*, *mixed-achievement grouping*, and *student-centered instruction* were each unrelated to achievement gains. The alternative regression analysis that included all of the teacher practice variables simultaneously did not substantively affect any of the positive relationships (see the last column in table 1). However, the comprehension variable, which had a null relationship to achievement in the marginal model, showed a negative relationship in the full model, with an effect size of –0.05. Given high levels of correlation across some of the scales (see table B-5 in appendix B), and the difficulty of interpreting the coefficients in the full model, the reader is cautioned against interpreting this negative relationship as evidence that student achievement is harmed by an emphasis on comprehension.

Table 2 presents the results of the specification in which the achievement gain in reading was regressed on teacher characteristics rather than on instructional practice scale scores. In this specification, no relationships were found for teacher-reported certification,[12] coursework, or kindergarten teaching experience. Part-time employment status, however, was negatively associated with gains, with an effect size of –0.13.[13]

[11]The questionnaire included information about the number of days reading was taught each week and the number of minutes each day devoted to reading. The vast majority of teachers reported teaching at least some reading each day, and the cutoff of 90 minutes for the analyses described here was intended to provide a coarse indicator of exposure and to create roughly balanced group sizes.

[12]Readers are reminded that the nature of early childhood and elementary certification is variable across states; thus, results of the relationships between certification status and achievement gains exhibited in a nationally representative sample should be interpreted with caution.

[13]A sensitivity analysis was conducted that included all reading instructional practice scale scores in the regression in table 2. No change was found in the significance levels of the teacher qualifications.

Table 1. Standardized regression coefficients for the regression of kindergarten student achievement gains in reading on instructional practice scales: School year 1998–99

Characteristic	Instructional practice scales (standardized regression coefficients)							
	Analysis including phonics	Analysis including didactic instruction	Analysis including reading and writing skills	Analysis including reading and writing activities	Analysis including compre-hension	Analysis including mixed-achievement grouping	Analysis including student-centered instruction	Full analysis
Intercept	-0.32 *	-0.29 *	-0.30 *	-0.35 *	-0.36 *	-0.36 *	-0.36 *	-0.30*
Teacher-level variables								
90 minutes or more reading instruction per day	0.08 *	0.08 *	0.06 *	0.07 *	0.09 *	0.09 *	0.10 *	0.05
Full-day kindergarten	0.17 *	0.18 *	0.16 *	0.17 *	0.19 *	0.19 *	0.19 *	0.16*
Instructional practice scales								
Phonics	0.07 *	†	†	†	†	†	†	0.07*
Didactic instruction	†	0.09 *	†	†	†	†	†	0.06*
Reading and writing skills	†	†	0.10 *	†	†	†	†	0.09*
Reading and writing activities	†	†	†	0.05 *	†	†	†	0.05*
Comprehension	†	†	†	†	0.01	†	†	-0.05*
Mixed-achievement grouping	†	†	†	†	†	0.01	†	#
Student-centered instruction	†	†	†	†	†	†	-0.01	-0.07*
School-level variables								
School locale[1]								
Suburban	0.03	0.02	0.04	0.04	0.03	0.03	0.03	0.02
Rural	-0.04	-0.07	-0.02	-0.01	-0.04	-0.04	-0.05	-0.06
School control[2]								
Private	0.03	-0.04	0.01	0.04	0.02	0.02	#	-0.03
Enrollment[3]								
150-299 students	0.09	0.12	0.12	0.11	0.11	0.11	0.11	0.12
300-499 students	0.09	0.11	0.11	0.12	0.11	0.11	0.11	0.10
500-749 students	0.07	0.10	0.10	0.10	0.10	0.10	0.10	0.08
750 or more students	0.06	0.08	0.08	0.08	0.08	0.09	0.09	0.06
Percent minority enrollment	-0.05	-0.07	-0.09	-0.04	-0.04	-0.03	-0.03	-0.11
Percent free-lunch-eligible enrollment	#	-0.02	0.02	0.02	#	#	#	#
Region[4]								
West	0.20 *	0.17 *	0.19 *	0.20 *	0.20 *	0.20 *	0.20 *	0.16*
Midwest	0.08 *	0.06	0.08 *	0.08 *	0.07 *	0.07 *	0.07 *	0.07*
South	0.11 *	0.08 *	0.08 *	0.10 *	0.11 *	0.11 *	0.11 *	0.06

See notes at end of table.

Table 1. Standardized regression coefficients for the regression of kindergarten student achievement gains in reading on instructional practice scales: School year 1998–99—Continued

Characteristic	Instructional practice scales (standardized regression coefficients)							
	Analysis including phonics	Analysis including didactic instruction	Analysis including reading and writing skills	Analysis including reading and writing activities	Analysis including compre-hension	Analysis including mixed-achievement grouping	Analysis including student-centered instruction	Full analysis
Child-level variables								
Female	0.10 *	0.10 *	0.10 *	0.10 *	0.10 *	0.10 *	0.10 *	0.10*
Race/ethnicity								
Asian/Pacific Islander[5]	0.08	0.08	0.07	0.08	0.08	0.08	0.08	0.08
Black, non-Hispanic	-0.18 *	-0.19 *	-0.18 *	-0.18 *	-0.18 *	-0.18 *	-0.18 *	-0.18*
Hispanic	#	#	#	#	#	#	#	#
Other, non-Hispanic race	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
Socioeconomic status	0.05 *	0.05 *	0.05 *	0.05 *	0.05 *	0.05 *	0.05 *	0.05*
Age at fall assessment	#	#	#	#	#	#	#	#
Days between tests	0.17 *	0.16 *	0.16 *	0.16 *	0.17 *	0.17 *	0.17 *	0.16*
Kindergarten repeater	-0.33 *	-0.33 *	-0.34 *	-0.33 *	-0.33 *	-0.33 *	-0.33 *	-0.33*
Non-English primary home language	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.08

† Not applicable.
# Rounds to zero.
* $p < 05$.
[1]The reference group for school locale is urban.
[2]The reference group for school control is public.
[3]The reference group for school size is less than 150 students.
[4]The reference group for school region is Northeast.
[5]The reference group for race/ethnicity is White, non-Hispanic.
NOTE: Standard errors are presented in table E-1 of appendix E. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Table 2. Standardized regression coefficients for the regression of kindergarten student achievement gains in reading on teacher qualifications: School year 1998–99

Characteristic	Teacher characteristics (standardized regression coefficients)
Intercept	-0.24 *
Teacher-level variables	
90 minutes or more reading instruction per day	0.09 *
Full-day kindergarten	0.18 *
Years of kindergarten teaching experience	0.02
Experience squared	-0.01
Part-time employment status	-0.13 *
Full certification	-0.05
Master's degree	-0.01
Coursework in methods of teaching reading[1]	
1 course	-0.01
2 courses	-0.03
3 courses	-0.07
4 courses	-0.06
5 courses	-0.04
6 or more courses	-0.03
School-level variables	
School locale[2]	
Suburban	0.03
Rural	-0.04
School control[3]	
Private	0.01
Enrollment[4]	
150-299 students	0.10
300-499 students	0.11
500-749 students	0.10
750 or more students	0.08
Percent minority enrollment	-0.04
Percent free-lunch-eligible enrollment	-0.01
Region[5]	
West	0.20 *
Midwest	0.07 *
South	0.10 *

See notes at end of table.

Table 2. Standardized regression coefficients for the regression of kindergarten student achievement gains in reading on teacher qualifications: School year 1998–99—Continued

Characteristic	Teacher characteristics (standardized regression coefficients)
Child-level variables	
Female	0.10 *
Race/ethnicity[6]	
Asian/Pacific Islander	0.08
Black, non-Hispanic	-0.18 *
Hispanic	#
Other, non-Hispanic race	-0.03
Socioeconomic status	0.05 *
Age at fall assessment	#
Days between tests	0.17 *
Kindergarten repeater	-0.33 *
Non-English primary home language	0.06

Rounds to zero.
* $p < 05$.
[1]Courses are in methods of teaching reading; the reference group is zero courses.
[2]The reference group for school locale is urban.
[3]The reference group for school control is public.
[4]The reference group for school size is less than 150 students.
[5]The reference group for school region is Northeast.
[6]The reference group for race/ethnicity is White, non-Hispanic.
NOTE: Standard errors are presented in table E-2 of appendix E. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

School-Level Variables: As shown in tables 1 and 2, kindergartners' reading gains were related to the region in which the school was located. Students in the West and South showed higher gains than those in the Northeast, with effect sizes ranging from 0.16 to 0.20 in the West and from 0.08 to 0.11 in the South. Students in the Midwest also showed higher gains in all specifications except those that included the *didactic* scale (effect sizes range from 0.07 to 0.08). No other school-level characteristics were associated with achievement gains.

Child-Level Variables: Several characteristics of the children were related to reading achievement gains. Girls showed higher gains than boys, with an effect size of 0.10 for all scales. Black kindergartners showed smaller gains than White kindergartners (effect sizes range from –0.18 to –0.19).[14] Socioeconomic status was positively associated with gains (effect size of 0.05 for all scales). Children who were repeating kindergarten showed gains that were lower than those of first-time kindergartners (effect sizes range from –0.33 to –0.34).

Instructional Practice Scale Regressions for Reading

The results of the teacher-level reading instructional practice regressions are presented in table 3. As explained in section 3, the instructional practice scales were based on the frequency with which teachers reported performing the instructional practices identified in the items that composed each scale.

Teacher-Level Variables: Teaching reading for 90 minutes or more per day was positively associated with the emphasis teachers placed on each of the reading instructional practice scales, with effect sizes ranging from 0.13 to 0.35. This was to be expected since the scales were constructed on the basis of teacher reports of the frequency with which they used certain practices. In addition, and possibly for related reasons, the full-day kindergarten indicator was also positive in all of the reading scale regressions except those relating to *phonics* and *comprehension* (effect sizes range from 0.16 to 0.32).

Coursework in methods of teaching reading was associated with several of the instructional practice scales, with effect sizes ranging from 0.33 to 0.66. Teachers who reported taking two, three, five, or six or more courses in pedagogy related to reading reported placing a greater emphasis on *mixed-achievement grouping* than teachers who did not take any reading methods coursework. Taking four or six or more such courses was also associated with a greater emphasis on *student-centered instruction* than taking no reading methods coursework. Taking six or more courses was associated with a greater emphasis on *phonics* and *reading and writing activities* compared with no coursework. The amount of reading methods coursework was not related to

[14]Throughout this report, White refers to White, non-Hispanic; Black refers to Black, non-Hispanic; and other, non-Hispanic refers to American Indians, Alaska Natives, Pacific Islanders, and multiracial kindergartners. The reader is cautioned against interpreting the race coefficients as pure "race effects," since they may be picking up family-related factors correlated with race. This report focuses on the relationships between teacher-level variables and student achievement; thus, the child-level variables serve only as controls and should not be overinterpreted. Studies that focus exclusively on the Black-White test score gap and study achievement levels rather than gains, such as Fryer and Levitt (2004), modify their model specifications to adjust for many of the types of family-related variables that are correlated with race. Thus, they are able to erase the Black-White test score gap in some specifications using ECLS-K.

the frequency with which teachers used *didactic*, *comprehension*, and *reading and writing skills* activities.

Certification[15] and possession of a master's degree were unrelated to teachers' emphases on any of the instructional practice scales. Kindergarten teaching experience was negatively associated with the reported use of *student-centered instruction* (effect size of –0.09). Part-time status was negatively related to the emphasis placed on *phonics*, *didactic instruction*, *comprehension*, and *reading and writing activities* (effect sizes range from –0.23 to –0.52).

School-Level Variables: School-level characteristics played a role in several reading instructional practice scale regressions, although the relationships varied depending upon the scale. Teachers in private schools reported less emphasis on *phonics, comprehension, reading and writing activities,* and *mixed-achievement grouping* than public school teachers (effect sizes range from –0.26 to –0.63) and greater emphasis on *didactic instruction* (effect size of 0.49).

Rural teachers reported greater emphasis than urban teachers on a *didactic* approach (effect size of 0.31) but less emphasis on *comprehension*, *reading and writing activities*, *reading and writing skills*, and *student-centered instruction* (effect sizes range from –0.22 to –0.41). Suburban teachers reported less emphasis than urban teachers on *student-centered instruction*.

School size was unrelated to all scales except *phonics*, in which teachers in schools with 500–749 students reported a greater emphasis than those in schools with fewer than 150 students (effect size of 0.44). The percentage of minority students was positively associated with the frequency of instruction on the *phonics*, *didactic*, and *reading and writing skills* scales, with effect sizes of 0.42, 0.51, and 0.50, respectively. The percentage of free-lunch-eligible students was negatively associated with the frequency of instruction on the *reading and writing activities* scale (effect size of –0.34).

Regional variables were related to two scales: teachers in the West and South reported a greater emphasis than those in the Northeast on a *didactic* approach (effect sizes of 0.28 and 0.32, respectively), and teachers in the West reported a lower emphasis than those in the Northeast on *comprehension* (effect size of –0.21).

[15]Readers are reminded that the nature of early childhood and elementary certification is variable across states; thus, results of the relationships between certification status and instructional practices exhibited in a nationally representative sample should be interpreted with caution.

Table 3. Standardized regression coefficients for the regressions of kindergarten teachers' reading instructional practice scales on selected teacher qualifications and school characteristics: School year 1998–99

Characteristic	Instructional practice scales (standardized regression coefficients)						
	Phonics	Didactic instruction	Compre-hension	Reading and writing activities	Reading and writing skills	Student-centered instruction	Mixed-achievement grouping
Intercept	-1.08 *	-0.69 *	-0.41 *	-0.38	-0.58 *	-0.47 *	-0.43 *
Teacher-level variables							
90 minutes or more reading instruction per day	0.22 *	0.13 *	0.35 *	0.31 *	0.31 *	0.22 *	0.30 *
Full-day kindergarten	0.14	0.17 *	0.08	0.29 *	0.18 *	0.32 *	0.16 *
Part-time employment status	-0.52 *	-0.24 *	-0.23 *	-0.24 *	-0.09	-0.06	-0.21
Years of kindergarten teaching experience	0.07	-0.01	0.01	-0.04	-0.03	-0.09 *	0.01
Experience squared	-0.04	0.01	0.01	-0.02	-0.01	0.03	#
Full certification	-0.01	-0.01	0.02	0.03	0.03	0.10	0.02
Master's degree	-0.11	0.02	0.03	0.01	-0.03	0.06	0.07
Coursework in methods of teaching reading[1]							
1 course	0.30	-0.01	0.02	0.14	-0.02	0.08	0.29
2 courses	0.45	0.04	0.06	0.19	-0.03	0.29	0.35 *
3 courses	0.50	0.04	0.14	0.25	#	0.27	0.36 *
4 courses	0.59	0.16	0.23	0.29	0.16	0.42 *	0.30
5 courses	0.51	0.08	0.07	0.29	-0.01	0.33	0.39 *
6 or more courses	0.66 *	0.15	0.24	0.33 *	0.17	0.43 *	0.38 *
School-level variables							
School locale[2]							
Suburban	0.07	0.05	-0.12	-0.12	-0.09	-0.20 *	-0.11
Rural	0.11	0.31 *	-0.34 *	-0.41 *	-0.22 *	-0.38 *	-0.08
School control[3]							
Private	-0.26 *	0.49 *	-0.29 *	-0.63 *	0.03	-0.42 *	-0.35 *
Enrollment[4]							
150-299 students	0.39	-0.09	0.23	-0.01	-0.07	0.02	-0.16
300-499 students	0.41	-0.11	0.19	-0.07	#	-0.04	-0.11
500-749 students	0.44 *	-0.12	0.25	-0.08	0.05	-0.06	-0.06
750 or more students	0.41	-0.02	0.18	0.01	0.09	0.01	-0.24
Percent minority enrollment	0.42 *	0.51 *	0.20	0.11	0.50 *	0.18	-0.09
Percent free-lunch-eligible enrollment	-0.18	0.05	-0.25	-0.34 *	-0.17	-0.20	-0.07
Region[5]							
West	-0.23	0.28 *	-0.21 *	0.11	0.07	-0.14	0.06
Midwest	0.03	0.14	-0.12	-0.04	-0.11	-0.14	-0.05
South	0.01	0.32 *	0.01	0.12	0.27	-0.06	0.15

Rounds to zero.

* $p < .05$.

[1]Courses are in methods of teaching reading; the reference group is zero courses.

[2]The reference group for school locale is urban.

[3]The reference group for school control is public.

[4]The reference group for school size is less than 150 students.

[5]The reference group for school region is Northeast.

NOTE: Standard errors are presented in table E-3 of appendix E. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Mathematics Achievement Gain Regressions

The regressions of gains in mathematics achievement also showed a number of relationships with teacher-, school-, and child-level characteristics. The regression results are presented in tables 4 and 5.

Teacher-Level Variables: As was the case with reading achievement gains, kindergartners' mathematics gains were related to the amount of time spent in school and on subject as well as on various teacher-reported instructional approaches but were unrelated to the qualifications of their teachers. Students in full-day kindergarten programs and students whose teachers reported spending more time on subject—in this case, 60 minutes or more per day teaching mathematics[16]—exhibited larger gains in mathematics achievement over the course of the kindergarten year than those who reportedly received less instruction, with effect sizes ranging from 0.06 to 0.10 for mathematics instruction time and from 0.14 to 0.18 for full-day kindergarten (table 4). Regarding the effects of different instructional practices, students of teachers who placed a greater emphasis on *traditional practices and computation* achieved greater gains than those whose teachers placed less emphasis on such practices (effect size of 0.11). The *advanced numbers and operations* and *measurement and advanced topics* scales were also positively associated with achievement gains—the effect sizes for these were 0.05 and 0.06, respectively. Students of teachers who placed a greater emphasis on *student-centered instruction* also showed higher gains than those who placed less emphasis on such practices (effect size of 0.03). As with reading, the alternative specification that included all of the teacher practice variables simultaneously did not substantively affect any of the positive relationships (see the last column of table 4). However, the numbers and geometry variable, which had a null relationship to achievement in the marginal model, showed a negative relationship in the full model (effect size of –0.06). Again, given correlations across scales (see table B-6 in appendix B), and the difficulty of interpreting the coefficients in the full model, the reader is cautioned against interpreting this as evidence that student achievement is harmed by an emphasis on numbers and geometry.

Table 5 displays the coefficients in the regression of mathematics achievement gains on teacher-reported qualifications rather than instructional practice scale scores. No statistically significant relationships were found between achievement and teachers' certification, coursework, kindergarten teaching experience, or part-time employment status.[17]

School-Level Variables: Kindergartners' gains in mathematics were related to region, school size, the percentage of minority students, and school type (public or private). As was found in the reading achievement gain regressions, students in the West, Midwest, and South showed greater mathematics gains than those in the Northeast (effect sizes range from 0.10 to 0.16) (table 4). In contrast to the results for reading gains, which showed no relationship to school size, students in schools with 500 or more students showed larger mathematics achievement gains than those in

[16]A cut score of 60 minutes was chosen for mathematics (in contrast to the 90 minutes used for reading) because teachers reported spending less time on mathematics than on reading, and a cut score of 60 created a roughly balanced pair of groups.

[17]A sensitivity analysis was conducted that included all mathematics instructional practice scale scores in the regression in table 5. No change was found in the significance levels of the teacher qualifications.

very small schools (i.e., those with fewer than 150 students) in all of the specified models (effect sizes range from 0.12 to 0.13). In addition, the mathematics achievement gains of children attending schools with 150–299 students were greater than those made by children in schools with fewer students in the specifications for the *advanced numbers and operations* and *traditional practices and computation* scales (effect sizes range from 0.11 to 0.12). Unlike the findings for reading achievement (table 1), the percentage of minority students was negatively associated with mathematics achievements gains in all of the specifications, with effect sizes ranging from –0.13 to –0.17 (table 4). In addition, students in private schools showed smaller mathematics achievement gains than those in public schools in the specifications for the *traditional practices and computation* scales (effect size of –0.08).

Child-Level Variables: Several child-level characteristics were related to mathematics achievement gains, although these relationships were, in some cases, different from those obtained in the reading achievement regressions. Both Black and Hispanic kindergartners showed smaller gains than White kindergartners (effect sizes range from –0.24 to –0.25 for Black children and were –0.07 for Hispanic children for all scales). In the case of Hispanic kindergartners, the negative association with achievement gains was partially driven by the set of children who did not reach the cutoff score on the OLDS screening test. When children who took at least one of the two mathematics tests in Spanish were removed from the sample (798 out of the 16,284), the Hispanic coefficient was no longer significant in three of the specifications (those containing the *traditional* and the *numbers and geometry* scales and the specification containing qualifications) (data not shown in tables). Removing this group of children did not affect coefficients on any of the teacher variables.

As was the case with the reading achievement gain regressions, children living in families with higher socioeconomic status exhibited greater gains than children from families with lower socioeconomic status (effect size of 0.06 for all scales), children who were repeating kindergarten showed smaller gains than first-time kindergartners (effect sizes range from –0.21 to –0.22), and a larger time gap between the fall and spring tests was positively associated with gains (effect size of 0.13) for all scales.

Table 4. Standardized regression coefficients for the regressions of kindergarten student achievement gains in mathematics on instructional practice scales: School year 1998–99

Characteristic	Instructional practice scales (standardized regression coefficients)						
	Analysis including numbers and geometry	Analysis including advanced numbers and operations	Analysis including traditional practices and computation	Analysis including measurement and advanced topics	Analysis including mixed-achievement grouping	Analysis including student-centered instruction	Full analysis
Intercept	-0.28 *	-0.27 *	-0.19 *	-0.27 *	-0.28 *	-0.27 *	-0.19 *
Teacher-level variables							
60 minutes or more mathematics instruction per day	0.10 *	0.08 *	0.06 *	0.08 *	0.09 *	0.09 *	0.06 *
Full-day kindergarten	0.18 *	0.17 *	0.14 *	0.16 *	0.18 *	0.17 *	0.14 *
Instructional practice scales							
Numbers and geometry	#	†	†	†	†	†	-0.06 *
Advanced numbers and operations	†	0.05 *	†	†	†	†	0.02 *
Traditional practices and computation	†	†	0.11 *	†	†	†	0.11 *
Measurement and advanced topics	†	†	†	0.06 *	†	†	0.03 *
Mixed-achievement grouping	†	†	†	†	0.02	†	0.01
Student-centered instruction	†	†	†	†	†	0.03 *	#
School-level variables							
School locale[1]							
Suburban	0.03	0.04	0.02	0.04	0.03	0.04	0.02
Rural	0.03	0.04	#	0.04	0.03	0.03	0.01
School control[2]							
Private	-0.01	#	-0.08 *	#	#	#	-0.08 *
Enrollment[3]							
150-299 students	0.10	0.11 *	0.12 *	0.11	0.11	0.10	0.12 *
300-499 students	0.10	0.10	0.10 *	0.10	0.10	0.10	0.10 *
500-749 students	0.12 *	0.13 *	0.13 *	0.13 *	0.13 *	0.12 *	0.13 *
750 or more students	0.13 *	0.13 *	0.12 *	0.13 *	0.13 *	0.13 *	0.11 *
Percent minority enrollment	-0.13 *	-0.13 *	-0.17 *	-0.13 *	-0.13 *	-0.14 *	-0.15 *
Percent free-lunch-eligible enrollment	-0.01	0.01	-0.03	#	#	#	-0.03
Region[4]							
West	0.16 *	0.15 *	0.16 *	0.17 *	0.16 *	0.16 *	0.15 *
Midwest	0.14 *	0.13 *	0.14 *	0.15 *	0.14 *	0.14 *	0.13 *
South	0.13 *	0.11 *	0.11 *	0.13 *	0.12 *	0.12 *	0.10 *

See notes at end of table.

Table 4. Standardized regression coefficients for the regressions of kindergarten student achievement gains in mathematics on instructional practice scales: School year 1998–99—Continued

Characteristic	Instructional practice scales (standardized regression coefficients)						
	Analysis including numbers and geometry	Analysis including advanced numbers and operations	Analysis including traditional practices and computation	Analysis including measurement and advanced topics	Analysis including mixed-achievement grouping	Analysis including student-centered instruction	Full analysis
Child-level variables							
Female	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
Race/ethnicity[5]							
Asian/Pacific Islander	0.02	0.01	0.02	0.02	0.02	0.02	0.02
Black, non-Hispanic	-0.24 *	-0.24 *	-0.25 *	-0.24 *	-0.24 *	-0.24 *	-0.25 *
Hispanic	-0.07 *	-0.07 *	-0.07 *	-0.07 *	-0.07 *	-0.07 *	-0.07 *
Other, non-Hispanic race	-0.03	-0.03	-0.02	-0.03	-0.03	-0.03	-0.02
Socioeconomic status	0.06 *	0.06 *	0.06 *	0.06 *	0.06 *	0.06 *	0.06 *
Age at fall assessment	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Days between tests	0.13 *	0.13 *	0.13 *	0.13 *	0.13 *	0.13 *	0.13 *
Kindergarten repeater	-0.22 *	-0.22 *	-0.21 *	-0.22 *	-0.22 *	-0.22 *	-0.21 *
Non-English primary home language	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02

† Not applicable.
# Rounds to zero.
* $p < .05$.
[1]The reference group for school locale is urban.
[2]The reference group for school control is public.
[3]The reference group for school size is less than 150 students.
[4]The reference group for school region is Northeast.
[5]The reference group for race/ethnicity is White, non-Hispanic.
NOTE: Standard errors are presented in table E-4 of appendix E. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Table 5. Standardized regression coefficients for the regression of kindergarten student achievement gains in mathematics on teacher qualifications: School year 1998–99

Characteristic	Teacher characteristics (standardized regression coefficients)
Intercept	-0.22 *
Teacher-level variables	
60 minutes or more mathematics instruction per day	0.10 *
Full-day kindergarten	0.18 *
Years of kindergarten teaching experience	0.01
Experience squared	-0.01
Part-time employment status	-0.02
Full certification	-0.05
Master's degree	0.01
Coursework in methods of teaching mathematics[1]	
1 course	0.01
2 courses	#
3 courses	-0.04
4 courses	-0.04
5 courses	-0.03
6 or more courses	0.04
School-level variables	
School locale[2]	
Suburban	0.03
Rural	0.03
School control[3]	
Private	-0.02
Enrollment[4]	
150-299 students	0.11
300-499 students	0.10
500-749 students	0.13 *
750 or more students	0.13 *
Percent minority students	-0.13 *
Percent free-lunch-eligible	-0.01
Region[5]	
West	0.16 *
Midwest	0.14 *
South	0.13 *

See notes at end of table.

Table 5. Standardized regression coefficients for the regression of kindergarten student achievement gains in mathematics on teacher qualifications: School year 1998–99—Continued

Characteristic	Teacher characteristics (standardized regression coefficients)
Child-level variables	
Female	-0.02
Race/ethnicity[6]	
Asian/Pacific Islander	0.02
Black, non-Hispanic	-0.24 *
Hispanic	-0.07 *
Other, non-Hispanic race	-0.02
Socioeconomic status	0.06 *
Age at fall assessment	0.01
Days between tests	0.13 *
Kindergarten repeater	-0.22 *
Non-English primary home language	-0.03

Rounds to zero.
* $p < .05$.
[1]Courses are in methods of teaching mathematics; the omitted category is zero courses.
[2]The reference group for school locale is urban.
[3]The reference group for school control is public.
[4]The reference group for school size is less than 150 students.
[5]The reference group for school region is Northeast.
[6]The reference group for race/ethnicity is White, non-Hispanic.
NOTE: Standard errors are presented in table E-5 of appendix E. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Instructional Practice Scale Regressions for Mathematics

The results for the mathematics instructional practice regressions are displayed in table 6. As explained in section 3, the instructional practice scales were based on the frequency with which teachers reported performing the instructional practices identified in the items that compose each scale.

Teacher-Level Variables: Teaching mathematics for 60 minutes or more per day was positively associated with the emphasis teachers placed on each of the mathematics instructional practice scales, as expected, with effect sizes ranging from 0.23 to 0.30. In addition, the full-day indicator was also positive in all of the mathematics instructional practice scale regressions except that relating to *advanced numbers and operations* (effect sizes range from 0.15 to 0.30).

The amount of coursework that teachers reported taking in methods of teaching mathematics was associated with scores on all instructional practice scales except the *measurement and advanced topics* scale, with effect sizes ranging from 0.08 to 0.37. Teachers who reported taking three, four, or six or more courses in pedagogy related to mathematics placed a greater emphasis on *number and geometry* and *advanced numbers and operations* than teachers who did not take any mathematics methods coursework. Taking four or six or more such courses was also associated with a greater emphasis on *traditional practices and computation* compared with no coursework. Taking three or more courses in mathematics methods was associated with a greater emphasis on *student-centered instruction* compared with no such coursework. Additionally, teachers who took two or more mathematics methods courses placed a greater emphasis on *mixed-achievement grouping* than teachers with no such coursework.

Kindergartners whose teachers reported that they were fully certified participated more frequently in *measurement and advanced topics* than those whose teachers were not fully certified (effect size of 0.23). The possession of a master's degree was not related to any scale score. Kindergarten teaching experience was positively associated with the reported frequency of the use of *mixed-achievement grouping,* with an effect size of 0.04. Part-time status was negatively associated with the emphasis placed on the *traditional practices and computation* scale (effect size of –0.26).

School-Level Variables: Certain school-level characteristics were related to the frequency with which teachers used particular instructional practices. Compared with teachers in public schools, teachers in private schools placed more emphasis on *traditional practices and computation* and less emphasis on *mixed-achievement grouping*, with effect sizes of 0.51 and –0.26, respectively. Teachers in rural schools placed less emphasis than those in urban schools on *student-centered instruction* and *measurement and advanced topics* (effect sizes of –0.25 and –0.26, respectively).

School size was unrelated to any of the mathematics instructional practice scale scores. The higher the percentage of minority students in the school, the greater the teachers' emphasis on *numbers and geometry*, *traditional practices and computation*, and *student-centered instruction* (effect sizes of 0.27, 0.37, and 0.30, respectively).

Regional differences varied. Teachers in the West reported spending more time than those in the Northeast on *advanced numbers and operations* but less time on *numbers and geometry* and

measurement and advanced topics (effect sizes of 0.20, –0.22, and –0.27, respectively). Teachers in the South and in the Midwest reported spending more time than those in the Northeast on *advanced numbers and operations* (effect sizes of 0.34 and 0.22, respectively). Teachers in the South also reported spending more time than those in the Northeast on *mixed-achievement grouping,* with an effect size of 0.19.

Sensitivity Analyses

In addition to the specifications presented above, a number of other specifications were used to test the sensitivity of results, particularly with respect to teacher certification. Teachers with temporary certification are often only a few months from earning full certification and may thus have similar levels of preservice training. A sensitivity analysis was run using a specification that substituted the full certification indicator with a variable indicating whether a teacher held a full or temporary certification in the achievement regressions. This variable behaved no differently than the full certification indicator. The achievement regressions that included teacher certification were also rerun without the variables indicating the number of courses completed in pedagogy to see if certification might then show a significant effect. The exclusion of these variables did not result in a change of significance for certification.

Table 6. Standardized regression coefficients for the regressions of kindergarten teachers' mathematics instructional practice scales on selected teacher qualifications and school characteristics: School year 1998–99

Characteristic	Instructional practice scales (standardized regression coefficients)					
	Numbers and geometry	Advanced numbers and operations	Traditional practices and computation	Student-centered instruction	Mixed-achievement grouping	Measurement and advanced topics
Intercept	-0.27	-0.19	-0.67 *	-0.62 *	-0.43 *	-0.36 *
Teacher-level variables						
60 minutes or more mathematics instruction per day	0.25 *	0.24 *	0.28 *	0.30 *	0.23 *	0.29 *
Full-day kindergarten	0.15 *	0.04	0.30 *	0.19 *	0.24 *	0.25 *
Part-time employment status	-0.16	-0.16	-0.26 *	-0.12	-0.12	0.05
Years of kindergarten teaching experience	#	-0.03	-0.03	#	0.04 *	0.02
Experience squared	0.01	-0.02	#	0.01	-0.01	-0.02
Full certification	0.03	-0.10	-0.10	-0.01	#	0.23 *
Master's degree	-0.03	-0.03	-0.04	0.04	0.02	-0.01
Coursework in methods of teaching mathematics[1]						
1 course	0.03	0.17	0.04	0.08	0.10	0.04
2 courses	0.14	0.13	0.01	0.14	0.08 *	0.07
3 courses	0.27 *	0.22 *	0.17	0.33 *	0.17 *	0.16
4 courses	0.40 *	0.25 *	0.25 *	0.31 *	0.18 *	0.17
5 courses	0.25	0.17	0.10	0.41 *	0.28 *	0.15
6 or more courses	0.37 *	0.29 *	0.27 *	0.35 *	0.16 *	0.18
School-level variables						
School locale[2]						
Suburban	-0.10	-0.01	0.09	-0.10	-0.04	-0.08
Rural	-0.16	-0.12	0.15	-0.25 *	-0.06	-0.26 *
School control[3]						
Private	-0.15	-0.16	0.51 *	-0.18	-0.26 *	-0.15
Enrollment[4]						
150-299 students	-0.24	-0.12	-0.18	0.16	-0.16	-0.08
300-499 students	-0.19	-0.21	-0.20	0.07	-0.08	-0.07
500-749 students	-0.14	-0.07	-0.16	0.12	-0.03	-0.11
750 or more students	-0.13	-0.06	-0.02	0.08	-0.08	-0.02
Percent minority enrollment	0.27 *	-0.05	0.37 *	0.30 *	0.10	0.05
Percent free-lunch-eligible enrollment	-0.10	-0.28	0.33	-0.22	#	-0.05
Region[5]						
West	-0.22 *	0.20 *	0.10	-0.05	0.09	-0.27 *
Midwest	-0.11	0.22 *	0.02	-0.06	0.03	-0.14
South	0.09	0.34 *	0.18	0.17	0.19 *	0.01

Rounds to zero.
* $p < .05$.
[1]Courses are in methods of teaching mathematics; the reference group is zero courses.
[2]The reference group for school locale is urban.
[3]The reference group for school control is public.
[4]The reference group for school size is less than 150 students.
[5]The reference group for school region is Northeast.
NOTE: Standard errors are presented in table E-6 of appendix E. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

6. Summary

This study contributes to the research investigating the importance of teachers in promoting student achievement. The analyses presented examine the relationship between student achievement in the kindergarten year and the self-reported qualifications of teachers, as has been done in several studies of older populations of children. The analyses also make use of self-reported information on specific teaching practices provided in ECLS-K survey data to extend the approach generally followed in the literature by looking at the association of these instructional practices with achievement. In addition, this study investigates relationships between the propensity to use these practices and teacher qualifications.

Limitations

Several caveats apply in interpreting the findings of this study. Many of these have been noted throughout the preceding sections but are worth reiterating here. Despite the many strengths of ECLS-K data for addressing research questions about teachers' effects on students, efforts to generalize from this study are hindered by limitations that are common to survey research.

Understanding the factors that contribute to effective teaching is limited partially as a result of incomplete measurement of those factors. The data on instructional practices provide some evidence regarding the frequency of particular approaches, for example, but do not address variation in the quality of implementation of those practices or how they may change over the course of the year. Furthermore, the instructional practice scales are based on self-reported information, and teachers' perceptions of what they do may not always match what they actually do in the classroom. Other data collection methods, such as classroom observations or teacher time logs, may yield more precise data about content coverage and instructional techniques at specific time periods than one-time teacher questionnaires; however, they are more costly to administer than questionnaires and they require much more frequent data collection to produce reliable estimates of instruction over the entire academic year (Rowan, Camburn, and Correnti 2004). While teacher questionnaires administered on an annual basis are less expensive than observation or time log data collections, they are limited by concerns of teachers' accuracy in recalling the curriculum topics emphasized over the entire school year. Furthermore, teachers may vary in the estimation techniques they use to report retrospectively on instructional activities over the entire year, resulting in different reports for teachers with similar teaching behaviors. However, by grouping multiple items that measure a common underlying characteristic, as is done in this study, the reliability of the composite variable will always be higher than the reliability of the individual items (Carmines and Zeller 1979).

Some important aspects of instruction, such as the cognitive complexity of the work teachers assign and the kind of feedback teachers provide on written work, are especially difficult to measure with paper-and-pencil surveys. ECLS-K and all surveys are limited in their ability to adequately capture teaching practices information, as each of the practice types may encompass many different teaching activities. For example, an item such as *doing an activity or project related to a book or story* could be interpreted and rated differently by teachers. Moreover, there is no information on the specific curriculum teachers are using. It is also important to keep in

mind that even if the scales used in this study accurately capture the instructional practice constructs they are designed to measure, the relationship between them and student achievement may not be fully captured in these assessments. For example, the assessments may not address all aspects of learning that are related to the instructional practices discussed. In addition, the instructional practices may affect later learning in ways that are not measurable in the kindergarten assessments. For example, teaching reading comprehension in kindergarten may prove crucial to future reading performance as well as science and analytic performance in later grades. The reader is therefore cautioned against interpreting the results of this study as clear signals for the adoption of policies that promote one instructional approach over another.

In addition, reliability estimates for certain instructional practice scales were low due to the small numbers of items comprising these scales. In particular, the coefficient alpha reliability estimate for the *didactic instruction* scale in reading (which consisted of three items) was 0.50, and the alpha estimates for the *mixed-achievement grouping* scales in reading and mathematics (both of which consisted of two items) were 0.48 and 0.56, respectively. Therefore, the results obtained using these scales should be interpreted with a greater degree of caution than those that used the other scales, which had reliability estimates ranging from approximately 0.7 to approximately 0.8.[18]

Measurement issues also arise when attempting to quantify or characterize teacher training. Teacher certification indicators may not measure consistent levels of preservice preparation, as this preparation varies from state to state. These inconsistencies may weaken the observable effects of these variables. Information on the numbers of relevant courses teachers completed during their preservice training is useful, but undoubtedly the quality and focus of those courses varies and there is no way to capture this variation in the models estimated in this study.

Limitations in the outcome measure should also be considered. The student assessments may accurately evaluate particular elements of reading and mathematics learning but leave other elements unmeasured or less well measured. A related point is that because of the exclusion rules implemented in this study and the methods used for assessing students whose primary language is not English, the findings do not necessarily generalize to English language learners.

A final caveat is that, since neither students nor teachers are randomly assigned to schools and students may not be randomly assigned to teachers, the analyses in this study may be subject to selection bias. This is a limitation shared by all studies using nonexperimental survey data. Although the presence of many control variables strengthens the model specifications used in this study, there may still be unobservable factors that systematically influence the matching of children to teachers and schools. It is therefore inadvisable to infer causal links between the teacher- and school-level variables examined and student outcomes. Despite these limitations, the rich data, their nested structure, and the longitudinal nature of the assessments in ECLS-K permit analyses that provide new information regarding existing relationships between student achievement, instructional practices, and teacher qualifications for the kindergarten population.

---

[18]Adjusting the regression coefficients for attenuation due to low reliability would result in stronger measured relationships than are reported here.

Discussion

The first research question addressed in this study asks how specific instructional approaches and teacher characteristics are related to achievement gains. With regard to this question, it was found that several teacher-reported variables describing instructional practices and organization were related to test-score gains. Among those factors that were more organizational than pedagogical, the results indicate that spending more time on subject and working within a full-day kindergarten structure were associated with relatively large gains in achievement compared with spending less time on subject or working in a part-day kindergarten setting. Among the instructional practices that teachers reported on in this study, those that emphasized reading and writing skills, didactic instruction, phonics, and reading and writing activities were positively associated with reading achievement gains. Instructional emphases on traditional practices and computation, measurement and advanced topics, advanced numbers and operations, and student-centered instruction (e.g., having students explain how problems were solved) were positively associated with mathematics achievement gains. The study provided no evidence of a direct relationship between the self-reported qualifications of teachers and student achievement with the exception of teachers' employment status. Kindergartners whose teachers were employed part time made smaller reading achievement gains than those whose teachers were employed full time.

The analyses conducted in response to the second research question (i.e., the relationship between teacher qualifications and the use of specific instructional practices) found evidence that certain teacher-reported background variables were positively associated with the use of various practices that, in turn, were associated with higher achievement. In particular, the number of courses teachers reported taking in methods of teaching reading and mathematics was related to the emphasis placed on certain instructional approaches. The completion of coursework in methods of teaching reading was positively associated with the use of phonics instruction, mixed-achievement grouping, student-centered instruction, and reading and writing activities. Coursework in methods of teaching mathematics was positively associated with the use of practices that emphasized numbers and geometry, advanced numbers and operations, traditional practices and computation, student-centered instruction, and mixed-achievement grouping. In addition, kindergarten teaching experience was negatively related to the use of student-centered instruction in reading and positively related to the use of mixed-achievement grouping in mathematics. Teacher certification appeared unrelated to reported instructional practices, with the exception of a positive association with an emphasis on measurement and advanced topics in mathematics.

Most of the effect sizes observed in this study are small. For example, the effect sizes for the instructional practice variables that showed significant relationships with achievement range from approximately 0.03 to 0.10. Although small, they are consistent with those found in other studies of relationships between instructional practice and achievement (e.g., Hamilton et al. 2003; Xue and Meisels 2004). Moreover, it should be kept in mind that these effect sizes represent relationships for only a single academic year. If the same relationships were observed in subsequent grade levels, the cumulative effect sizes would be larger and more meaningful. As discussed below, the findings suggest the need to examine these relationships over a longer period of time.

Future Research Using ECLS-K Data

This study focused on teacher qualifications, preservice training, and frequency-based measures of instructional practice as factors related to student achievement. Although these factors are likely to be important, ECLS-K data include several additional categories of information that may also be associated with student outcomes, including teachers' assessments of the quality of their schools' instructional climate, their philosophies about student learning, their evaluations of the adequacy of resources in their schools, and their levels of job satisfaction. Extensions of this study could incorporate these constructs. In addition, the data include information about teachers' ongoing professional development opportunities, which may be as important as their preservice education. Incorporating these into the models would provide a more complete picture of the kinds of training associated with positive student outcomes.

Another important feature of ECLS-K data is the tracking of students over the course of several years. The kinds of teacher qualifications and practices that are associated with positive kindergarten outcomes may differ from those that contribute to learning among older elementary students, so incorporating future waves of ECLS-K data would be informative. In addition, having multiple measures of student achievement throughout elementary school would facilitate the use of more complex longitudinal models, which would shed light on the nature of students' achievement growth and the ways in which it is associated with teachers' characteristics and classroom practices.

References

Adams, M.J. (1990). *Beginning to Read: Thinking and Learning About Print*. Cambridge, MA: Massachusetts Institute of Technology.

Ball, E.W., and Blachman, B.A. (1991). Does Phoneme Awareness Training in Kindergarten Make a Difference in Early Word Recognition and Developmental Spelling? *Reading Research Quarterly, 26*(1): 49–66.

Baroody, A.J. (1989). Manipulatives Don't Come With Guarantees. *The Arithmetic Teacher, 37:* 4–5.

Carmines, E.G., and Zeller, R.A. (1979). *Reliability and Validity Assessment*. Newbury Park, CA: Sage Publications.

Chall, J. (1992). The New Reading Debates: Evidence From Science, Art and Ideology. *Teachers College Record, 94*(2): 315–328.

Choi, N., Fuqua, D.R., and Griffin, B.W. (2001). Exploratory Analysis of the Structure of Scores From the Multidimensional Scales of Perceived Self-Efficacy. *Educational and Psychological Measurement, 61:* 475–489.

Cohen, D.K., and Hill, H.C. (2000). Instructional Policy and Classroom Performance: The Mathematics Reform in California. *Teachers College Record,102:* 294–343.

D'Agostino, J.D. (2000). Instructional and School Effects on Students' Longitudinal Reading and Mathematics Achievements. *School Effectiveness and School Improvement, 11*(2): 197–235.

Denton, K., and West, J. (2002). *Children's Reading and Mathematics Achievement in Kindergarten and First Grade* (NCES 2002-125). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Duncan, S.E., and De Avila, E.A. (1998). PreLAS 2000 Cue Picture Book English Form C. Monterey, CA: CTB/McGraw-Hill Companies.

Fetler, M. (1999). High School Characteristics and Mathematics Test Results. *Education Policy Analysis Archives, 7*(9).

Fryer, Jr., R., and 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.

Goldhaber, D., and Brewer, D. (2000). Does Teacher Certification Matter? High School Certification Status and Student Achievement. *Educational Evaluation and Policy Analysis, 22:* 129–145.

Gottman, J.M., and Rushe, R.H. (1993). The Analysis of Change: Issues, Fallacies and New Ideas. *Journal of Counseling and Clinical Psychology, 61:* 907–910.

Germino Hausken, E., Walston, J., and Rathbun, A. (2004). *Kindergarten Teachers: Public and Private School Teachers of the Kindergarten Class of 1998–99* (NCES 2004-060). Washington, DC: U.S. Government Printing Office.

Hamilton, L.S., McCaffrey, D., Klein, S.P., Stecher, B.M., Robyn, A., and Bugliari, D. (2003). Teaching Practices and Student Achievement: Studying Classroom-Based Education Reforms. *Educational Evaluation and Policy Analysis, 25:* 1–29.

Hamilton, L.S., and Guarino, C.M. (2005). *Measuring the Practices, Philosophies, and Characteristics of Kindergarten Teachers* (Working Paper WR-199-EDU). Santa Monica, CA: RAND Corporation.

Hawk, P., Coble, C., and Swanson, M. (1985). Certification: It Does Matter. *Journal of Teacher Education*, *36*(3): 13–15.

Maris, E. (1998). Covariance Adjustment Versus Gain Scores-Revisited. *Psychological Methods*, *3:* 309–327.

Mayer, R. (1998). Do New Teaching Standards Undermine Performance on Old Tests? *Educational Evaluation and Policy Analysis, 20:* 53–73.

Mayer, R. (1999). Measuring Instructional Practice: Can Policymakers Trust Survey Data? *Educational Evaluation and Policy Analysis, 21*: 29–45.

Monk, D. (1994). Subject Area Preparation of Secondary Mathematics and Science Teachers and Student Achievement. *Economics of Education Review, 12:* 125–145.

Murnane, R., and Phillips, B. (1981). What Do Effective Teachers of Inner-City Children Have in Common? *Social Science Research, 10:* 83–100.

National Center for Education Statistics (2001). *ECLS-K Base Year Restricted-Use Salary and Benefits File* (NCES 2001-014). Washington, DC: U.S. Government Printing Office.

National Research Council. (2001). *Adding It Up: Helping Children Learn Mathematics.* J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Raudenbush, S., and Bryk, A. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks: Sage Publications.

Rock, D.A., and Pollack, J.M. (2002). *Early Childhood Longitudinal Study—Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for Kindergarten Through First Grade* (NCES 2002-05). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Rogosa, D.R., and Willet, J.B. (1983). Demonstrating the Reliability of the Difference Score in the Measurement of Change. *Journal of Educational Measurement, 20:* 335–343.

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

Rowan, B., Camburn, E., and Correnti, R. (2004). Using Teacher Logs to Measure the Enacted Curriculum in Large-Scale Surveys: A Study of Literacy Teaching in 3rd Grade Classrooms. *Elementary School Journal, 105:* 75–102.

Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley and Sons.

Schafer, J.L. (1997). *Analysis of Incomplete Multivariate Data*. New York: Chapman and Hall.

Singer, J. (1998). Using SAS Proc Mixed to Fit Multi-Level Models, Hierarchical Models, and Individual Growth Models. *Journal of Educational and Behavioral Statistics, 24:* 323–355.

Snow, C., Burns, S., and Griffin, P. (Eds.). (1998). *Preventing Reading Difficulties in Young Children.* Washington, DC: National Academy Press.

Sowell, E.J. (1989). Effects of Manipulative Materials in Mathematics Instruction. *Journal for Research in Mathematics Education*, *20:* 498–505.

Stipek, D., and Byler, P. (2004). The Early Childhood Classroom Observation Measure. *Early Childhood Research Quarterly, 19*(3): 375–397.

Thompson, B., and Daniel, L.G. (1996). Factor Analytic Evidence for the Construct Validity of Scores: A Historical Overview and Some Guidelines. *Educational and Psychological Measurement, 56:* 197–208.

Tourangeau, K., Burke, J., Le, T., Wan, S., Weant, M., Brown, E., Vaden-Kiernan, N., Rinker, E., Dulaney, R., Ellingson, K., Barrett, B., Flores-Cervantes, I., Zill, N., Pollack, J., Rock, D., Atkins-Burnett, S., Meisels, S., Bose, J., West, J., Denton, K., Rathbun, A., and Walston, J. (2001). *ECLS-K Base-Year Public-Use Data Files and Electronic Codebook: User's Manual* (NCES 2001-029). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Williams, R.H., and Zimmerman, D.W. (1996). Are Simple Gain Scores Obsolete? *Applied Psychological Measurement, 20:* 59–69.

Wright, S.P., Horn, S.P., and Sanders, W.L. (1997). Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. *Journal of Personnel Evaluation in Education, 11:* 57–67.

Xue, Y., and Meisels, S.J. (2004). Early Literacy Instruction and Learning in Kindergarten: Evidence From the Early Childhood Longitudinal Study—Kindergarten Class of 1998–1999. *American Educational Research Journal, 41:* 191–229.

Appendix A:
ECLS-K Direct Cognitive Assessment Measures

This appendix summarizes the direct cognitive assessments administered to ECLS-K sample children in the fall and spring of kindergarten. For more information about the ECLS-K cognitive assessments, please see the *ECLS-K Base-Year Public-Use Data Files and Electronic Codebook: User's Manual* (Tourangeau et al. 2001) and the *Early Childhood Longitudinal Study—Kindergarten Class of 1998–99 (ECLS-K), Psychometric Report for Kindergarten Through First Grade* (Rock and Pollack 2002). In the fall of the base year, one-on-one child assessments were conducted with the sampled children. The assessment took approximately 50–70 minutes and was designed to provide data on the developmental status of children in the United States at the start of their formal schooling. The ECLS-K cognitive assessment battery consisted of questions in three subject areas: language and literacy, mathematical thinking, and general knowledge.

Instrument Development

During the ECLS-K cognitive test development, an initial review of commercial assessments indicated that there were no "off-the-shelf" tests that met the domain requirements of ECLS-K, were both individually administered and adaptive, or provided items that could be used to measure children's cognitive achievement longitudinally. The framework for ECLS-K drew from the National Assessment of Educational Progress (NAEP) fourth-grade test specifications. The NAEP assessment goals are similar to those of ECLS-K in that both projects assess cognitive skills typically emphasized in schools. For the grades in which the NAEP frameworks were inappropriate, ECLS-K solicited advice from early elementary school educators and curriculum specialists to articulate more suitable test specifications. The expertise of item writers from Educational Testing Service (ETS), elementary school curriculum specialists, and elementary school teachers was also used to develop new assessment items and select existing items to borrow or adapt, with permission, from published tests, including the Peabody Individual Achievement Test-Revised (PIAT-R), the Peabody Picture Vocabulary Test-Revised (PPVT-R), the Primary Test of Cognitive Skills (PTCS), the Test of Early Reading Ability (TERA-2), the Test of Early Mathematics Ability (TEMA-2), and the Woodcock-Johnson Tests of Achievement-Revised (WJ-R) (Rock and Pollack 2002). Across domains, test items were reviewed by elementary school curriculum and content area specialists for appropriateness of content and difficulty and for relevance to the assessment battery framework. In addition, items were reviewed for issues related to sensitivity to minority concerns.

The Language Screener

Prior to administering the cognitive assessment battery, a brief language-screening assessment was administered to those children identified from their school records (or by their teacher, if no school records were available) as coming from a language minority background (meaning that their primary home language was not English). The Oral Language Development Scale (OLDS) screening test was used to determine if a child was able to understand and respond to the ECLS-K cognitive assessment items in English. Children who passed the language screener received

the full ECLS-K direct assessment battery. Children who did not pass an established cut score on the language screener received a reduced version of the ECLS-K assessments. The OLDS, extrapolated from parts of the PreLAS 2000 Assessment (Duncan and De Avila 1998), measured children's listening comprehension, vocabulary, and ability to understand and produce language. Children who passed the language screener received the full English direct assessment. Certain components of the direct child assessment could also be conducted in Spanish. If a child did not pass the language screener but spoke Spanish, he or she was administered a Spanish translated form of the mathematics assessment. Children who did not pass the established cut score on the language screener and whose native language was not Spanish were excluded from the cognitive assessment; however, assessors collected physical measurements of these children's height and weight. Children who did not pass the OLDS in the fall of kindergarten were reassessed with the OLDS at subsequent data collection waves until they passed the OLDS language screener. In the fall of kindergarten, 1,567 children were not administered the English battery because of their performance on the OLDS. By the spring of first grade, this number had been reduced to 350.

Two-Stage Assessment Design

The direct cognitive assessment consisted of a set of two-stage assessments: a first-stage routing section for each subject area, followed by several alternative second-stage forms. The purpose of this adaptive assessment design was to maximize accuracy of measurement and minimize administration time. The same reading and mathematics sections, consisting of 12 to 20 items with a broad range of difficulty, were administered to all children. The routing section provided a rough estimate of each child's achievement level, so that a second-stage form with items of the appropriate difficulty for maximizing measurement accuracy could be selected. A child's performance on the routing section determined the second-stage form that was administered. The reading and mathematics assessments had low-, middle-, and high-difficulty second-stage options. The second-stage forms varied by level of difficulty so that a child would be administered questions appropriate to his or her current level of ability for each cognitive domain. The number of questions included in this assessment was limited in order to minimize the time and burden on the children.

The cognitive assessment included both multiple-choice and open-ended items. For ease of administration, questions of similar format were grouped together in order of increasing difficulty within each group. When the question format changed, practice items were used to introduce children to the new format. Assessments were shortened or discontinued if the administrator perceived that the child was uncomfortable or distressed about responding to the assessment items. If no response was given to 10 questions in a row, assessors entered a "refuse" code into the computer for the remainder of the items in that subject area, without reading the questions, until reaching the next subject area, where he or she resumed reading the questions. This procedure was used to give children who did not want to respond to questions in one subject area (e.g., reading) a chance to respond to questions in another subject area (e.g., math). Scores in each subject area were computed only if at least 10 questions were answered in the combined first and second stages.

Cognitive Components

The cognitive assessment focused on three general areas of competence: (1) language use and literacy (reading); (2) mathematics; and (3) knowledge of the social and physical world, referred to as "general knowledge." The assessment did not ask the children to write anything or to explain their reasoning; rather, they used pointing or verbal responses to complete the tasks. The data were collected using a computer-assisted personal interviewing (CAPI) methodology. The assessment included the use of a small easel with pictures, letters of the alphabet, words, short sentences, numbers, or number problems. A brief description of the reading and mathematics cognitive assessments follows (the general knowledge assessment scores were not used in this report).

Reading: The reading assessment included questions designed to measure basic skills (print familiarity, letter recognition, beginning and ending sounds, rhyming sounds, word recognition), vocabulary (receptive vocabulary), and comprehension (listening comprehension, words in context). Comprehension items were targeted to measure skills in initial understanding, developing interpretation, personal reflection, and demonstrating critical stance.

Mathematics: The mathematics assessment items were designed to measure skills in conceptual knowledge, procedural knowledge, and problem solving. Approximately one-half of the mathematics assessment consisted of questions on number sense and number properties and operations. The remainder of the assessment included questions in measurement; geometry and spatial sense; data analysis, statistics, and probability; and patterns, algebra, and functions. Each of the mathematics assessment forms contained several items for which manipulatives were available for children to use in solving the problems. Paper and pencil were also offered to the children to use for the appropriate parts of the assessment.

Reading and Mathematics Assessment Scores

Scores based on the full set of test items were calculated using item response theory (IRT) procedures. IRT made it possible to calculate scores that could be compared regardless of which second-stage form a child took. IRT uses the pattern of right, wrong, and omitted responses to the items actually administered in a test and the difficulty, discriminating ability, and "guess-ability" of each item to place each child on a continuous ability scale. The items in the routing test, plus a core set of items shared among the different second-stage forms, made it possible to establish a common scale. It is then possible to estimate the score the child would have achieved if all of the items in all of the test forms had been administered. IRT scoring makes possible longitudinal measurement of gain in achievement over time, even though the tests administered are not identical at each point. The common items present in the routing test and in overlapping second-stage forms allow the test scores to be placed on the same scale, even as the two-stage test design adapts to children's growth over time. Reliability statistics appropriate for each type of score were computed for each subject area, for the fall and spring kindergarten assessments. For the IRT-based scores, the reliability of the overall ability estimate, theta, is based on the variance of repeated estimates of theta. The fall and spring kindergarten reading and mathematics reliabilities range from 0.92 to 0.95. The IRT scale scores in the database represent estimates of the number of items students would have answered correctly if they had taken all of the 72 questions in the first- and second-stage reading forms and all of the 64 questions in the

mathematics forms. These scores are not integers because they are probabilities of correct answers, summed over all items in the pool. Gain scores may be obtained by subtracting the estimated number right at time 1 from the estimated number right at time 2.

Table A-1 shows the types of scores, range of values, means, and standard deviations for the direct cognitive assessment.

Table A-1. Means, standard deviations, and range of values for the ECLS-K base-year (kindergarten) reading and mathematics assessments: Fall 1998 and spring 1999

Assessment score	Mean	Standard deviation	Range of values
Fall kindergarten reading IRT scale score	22	8.3	0-72
Spring kindergarten reading IRT scale score	32	10.3	0-72
Fall kindergarten mathematics IRT scale score	19	7.2	0-64
Spring kindergarten mathematics IRT scale score	27	8.8	0-64

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Appendix B: Methodology Used in the Development of the Instructional Practice Scales

The development of instructional practice scales relies on prior work for the National Center for Education Statistics (NCES) in which factor analyses of the instructional practice items were conducted to inform the development of scales measuring specific pedagogical constructs (Hamilton and Guarino 2005). This appendix describes the main procedures used for constructing the scales; additional details, including discussions of sensitivity analyses, are available from the authors upon request.

Exploratory factor analysis was conducted to examine the structure of correlations among the instructional practice items in each subject. Each factor analysis was conducted on a random half sample of the data, which facilitated a cross-validation to test the stability of the solution. The initial analyses were conducted using one-half of the data. Once a reasonable solution was identified, the factor analysis was repeated using the second half of the data. The patterns of loadings across the two solutions were examined. The results were used to identify the number of items for which the highest loading occurred on different factors in the two half samples. In all cases, this number was small and was limited to those items that had relatively weak loadings on all factors in the original analysis. Thus, the cross-validation provided evidence regarding the stability of each solution.

All analyses used principal components analysis with orthogonal (varimax) rotation. Other factor extraction methods were explored, and most of the results were relatively robust to choice of method.[1] This finding is consistent with other work that has compared results from different extraction methods (Thompson and Daniel 1996). Because some of the factors were expected to be correlated, solutions using various oblique rotations were also examined. Conclusions about which items clustered together were consistent with those from the orthogonal rotation, a finding that is also consistent with other research (e.g., Choi, Fuqua, and Griffin 2001). This report focuses on the orthogonal solutions to simplify reporting of results. Decisions about the number of factors to retain in each case were made based on a combination of empirical considerations (e.g., scree plots) and the need for scales that would be instructionally meaningful and informative. All analyses were weighted using the B2TW0 weights.

Teachers who were missing data on one or more items were excluded from the initial factor analyses. A total of 3,243 teachers were administered the spring teacher questionnaires. The reading factor analysis was conducted on the 2,323 teachers who had complete data on all items. The mathematics analysis included 2,287 teachers.

[1]There were some exceptions in the reading analysis, discussed below.

To create scales for use in subsequent modeling, simple sum scores were calculated using the items that had their highest loading on a given factor.[2] This approach is consistent with the scale development used in other studies of instructional practice (see, e.g., Cohen and Hill 2000; Hamilton et al. 2003) and tends to produce results similar to those obtained through other methods of scale development. For teachers who answered some, but not all, items on a given scale, the average score of those items that they answered was used.

Table B-1 lists the scales, the corresponding items, and the weighted item means for reading. Table B-2 lists the scales for mathematics. Tables B-3 and B-4 present the internal consistency (coefficient alpha) reliabilities of the reading and mathematics items, respectively. As expected, the reliabilities of the didactic instruction scale in reading and the grouping scales in both subjects are low due to the small numbers of items on these scales. This suggests that these scales should be used with caution. Some of these scales may be combined in ways that would enhance reliability without sacrificing information. However, the scales with particularly low reliability in these analyses seem to address constructs that are distinct from what is tapped by the other scales, so for some purposes they may be useful. Tables B-5 and B-6 present the bivariate correlations between the instructional practice scales used in this report.

[2]For a justification for using sum scores and a discussion of alternative methods, including IRT, see Hamilton and Guarino (2005). In particular, the correlations between corresponding IRT and sum scores in this study were all greater than 0.90, and for three of the seven scales the correlations were 0.99.

Table B-1. Kindergarten teacher reading instructional practice scales and item means: Spring 1999

Instructional scale	Item number	Variable name	Description	Mean score	Standard deviation
Reading and writing skills	29f	A2SYLLAB	Reading multi-syllable words, like adventure	2.7	1.84
	29m	A2PNCTUA	Using capitalization and punctuation	4.1	1.91
	29n	A2COMPSE	Composing and writing complete sentences	3.5	1.98
	29o	A2WRTSTO	Composing and writing stories with an understandable beginning, middle, and end	2.2	1.57
	29p	A2SPELL	Conventional spelling	2.9	2.01
	29q	A2VOCAB	Vocabulary	4.3	1.94
	29r	A2ALPBTZ	Alphabetizing	2.1	1.61
	29s	A2RDFLNT	Reading aloud fluently	3.1	2.07
Phonics	28a	A2LRNLT	Work on learning the names of letters	5.8	0.70
	28b	A2PRACLT	Practice writing the letters of the alphabet	5.5	0.90
	28c	A2NEWVOC	Discuss new or difficult vocabulary	5.4	0.92
	28e	A2PHONIC	Work on phonics	5.7	0.77
	29a	A2CONVNT	Conventions of print	5.4	1.17
	29b	A2RCGNZE	Alphabet and letter recognition	5.8	0.70
	29c	A2MATCH	Matching letters to sounds	5.7	0.73
	29d	A2WRTNME	Writing own name (first and last)	5.5	1.15
	29e	A2RHYMNG	Rhyming words and word families	4.6	1.11
Didactic instruction	28j	A2BASAL	Read from basal reading texts	2.4	1.89
	28l	A2WRKBK	Work in a reading workbook or on a worksheet	4.1	1.84
	28m	A2WRTWRD	Write words from dictation, to improve spelling	2.9	1.80
Comprehension	29g	A2PREPOS	Common prepositions, such as over and under, up and down	3.9	1.57
	29h	A2MAINID	Identifying the main idea and parts of a story	4.1	1.68
	29i	A2PREDIC	Making predictions based on text	4.8	1.29
	29j	A2TEXTCU	Using context cues for comprehension	4.4	1.67
	29k	A2ORALID	Communicating complete ideas orally	5.3	1.15
	29l	A2DRCTNS	Remembering and following directions that include a series of actions	5.2	1.18
Student-centered reading instruction	28d	A2DICTAT	Dictate stories to a teacher, aide, or volunteer	3.9	1.37
	28h	A2RETELL	Retell stories	4.2	1.21
	28q	A2DOPROJ	Do an activity or project related to a book or story	4.0	1.35
	28r	A2PUBLSH	Publish their own writing	2.4	1.41
	28s	A2SKITS	Perform plays and skits	2.2	1.00
Reading and writing activities	28i	A2READLD	Read aloud	4.7	1.46
	28k	A2SILENT	Read silently	4.1	1.94
	28n	A2INVENT	Write with encouragement to use invented spellings, if needed	4.6	1.57
	28o	A2CHSBK	Read books they have chosen for themselves	5.0	1.34
	28p	A2COMPOS	Compose and write stories or reports	3.3	1.72
	28t	A2JRNL	Write stories in a journal	3.7	1.84
Mixed-achievement grouping	28v	A2MXDGRP	Work in mixed-achievement groups on language arts activities	4.4	1.79
	28w	A2PRTUTR	Peer tutoring	3.4	1.78

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, spring 1999.

Table B-2. Kindergarten teacher mathematics instructional practice scales and item means: Spring 1999

Instructional scale	Item number	Variable name	Description	Mean score	Standard deviation
Numbers and geometry	31a	A2OUTLOU	Count out loud	5.6	0.71
	31b	A2GEOMET	Work with geometric manipulatives	4.3	1.19
	31c	A2MANIPS	Work with counting manipulatives to learn basic operations	4.8	1.04
	32a	A2QUANTI	Correspondence between number and quantity	5.1	1.08
	32b	A21TO10	Writing all numbers between 1 and 10	4.6	1.29
	32f	A2SHAPES	Recognizing and naming geometric shapes	4.1	1.39
	32g	A2IDQNTY	Identifying relative quantity (e.g., equal, most, less, more)	4.3	1.26
	32h	A2SUBGRP	Sorting objects into subgroups according to a rule	3.9	1.24
	32i	A2SZORDR	Ordering objects by size or other properties	3.7	1.18
	32j	A2PTTRNS	Making, copying, or extending patterns	4.3	1.29
	32u	A2ORDINL	Ordinal numbers (e.g., first, second, third)	3.9	1.51
Advanced numbers and operations	32c	A22S5S10	Counting by 2s, 5s, and 10s	4.1	1.68
	32d	A2BYD100	Counting beyond 100	3.0	1.94
	32n	A2PLACE	Place value	3.0	2.13
	32o	A2TWODGT	Reading two-digit numbers	4.6	1.66
	32p	A23DGT	Reading three-digit numbers	2.7	2.02
Traditional practices and computation	31k	A2MTHSHT	Do mathematics worksheets	4.2	1.54
	31l	A2MTHTXT	Do mathematics problems from their textbooks	2.1	1.87
	31m	A2CHLKBD	Complete mathematics problems on the chalkboard	2.9	1.73
	32e	A2W12100	Writing all numbers between 1 and 100	2.3	1.55
	32k	A2REGZCN	Recognizing the value of coins and currency	3.4	1.50
	32l	A2SNGDGT	Adding single-digit numbers	4.0	1.48
	32m	A2SUBSDG	Subtracting single-digit numbers	3.5	1.63
Student-centered mathematics instruction	31d	A2MTHGME	Play mathematics-related games	4.5	1.14
	31f	A2MUSMTH	Use music to understand mathematics concepts	2.7	1.52
	31g	A2CRTIVE	Use creative movement or creative drama to understand mathematics concepts	2.6	1.39
	31i	A2EXPMTH	Explain how a mathematics problem is solved	3.9	1.58
	31n	A2PRTNRS	Solve mathematics problems in small groups or with a partner	3.5	1.52
	31o	A2REALLI	Work on mathematics problems that reflect real-life situations	3.9	1.46
Mixed-achievement grouping	31p	A2MXMATH	Work in mixed achievement groups on mathematics activities	4.1	1.81
	31q	A2PEER	Peer tutoring	3.1	1.80
Measurement and advanced topics	31h	A2RULERS	Work with rulers, measuring cups, spoons, or other measuring instruments	2.9	1.17
	32r	A2GRAPHS	Reading simple graphs	3.7	1.45
	32s	A2DATACO	Performing simple data collection and graphing	3.2	1.38
	32t	A2FRCTNS	Fractions (e.g., recognizing that 1/4 of a circle is colored)	2.0	1.17
	32v	A2ACCURA	Using measuring instruments accurately	2.6	2.56
	32w	A2TELLTI	Telling time	3.3	3.31
	32x	A2ESTQNT	Estimating quantities	3.0	1.36
	32bb	A2PRBTY	Estimating probability	1.7	1.15
	32cc	A2EQTN	Writing mathematics equations to solve word problems	1.9	1.36

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, spring 1999.

Table B-3. Reliabilities and correlations for kindergarten teacher reading instructional practice sum scores and IRT scores: Spring 1999

Instructional practices scale	Number of items	Alpha for sum score	Marginal reliability for IRT score	Correlation between sum and IRT score
Reading and writing skills	8	0.83	0.85	0.97
Phonics	9	0.75	0.68	0.90
Didactic instruction	3	0.50	0.53	0.99
Comprehension	6	0.77	0.82	0.94
Student-centered instruction	5	0.69	0.72	0.99
Reading and writing activities	6	0.77	0.82	0.96
Mixed-achievement grouping	2	0.48	0.51	0.99

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, spring 1999.

Table B-4. Coefficient alpha reliabilities for kindergarten teacher mathematics instructional practice scales: Spring 1999

Instructional practices scale	Number of items	Alpha for sum score
Numbers and geometry	11	0.83
Advanced numbers and operations	5	0.75
Traditional practices and computation	7	0.73
Student-centered mathematics instruction	6	0.74
Mixed-achievement grouping	2	0.56
Measurement and advanced topics	9	0.81

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, spring 1999.

Table B-5. Bivariate correlations of the kindergarten teacher reading instructional practice scales: Spring 1999

Instructional practices scale	Reading and writing skills	Phonics	Didactic instruction	Compre-hension	Student-centered instruction	Reading and writing activities	Mixed-achieve-ment grouping
Reading and writing skills	1.00	†	†	†	†	†	†
Phonics	.32	1.00	†	†	†	†	†
Didactic instruction	.35	.21	1.00	†	†	†	†
Comprehension	.52	.48	.17	1.00	†	†	†
Student-centered instruction	.41	.40	.11	.50	1.00	†	†
Reading and writing activities	.52	.37	.16	.48	.61	1.00	†
Mixed-achievement grouping	.30	.28	.11	.35	.37	.41	1.00

† Not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, spring 1999.

Table B-6. Bivariate correlations of kindergarten teacher mathematics instructional practice scales: Spring 1999

Instructional practices scale	Numbers and geometry	Advanced numbers and operations	Traditional practices and computation	Student-centered instruction	Mixed-achievement grouping	Measurement and advanced topics
Numbers and geometry	1.00	†	†	†	†	†
Advanced numbers and operations	.32	1.00	†	†	†	†
Traditional practices and computation	.36	.32	1.00	†	†	†
Student-centered instruction	.59	.29	.39	1.00		†
Mixed-achievement grouping	.37	.23	.23	.51	1.00	†
Measurement and advanced topics	.55	.43	.42	.56	.36	1.00

† Not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, spring 1999.

Appendix C:
Methodology Used in the Regression Analyses

Hierarchical linear modeling (HLM) was utilized to estimate the parameters in the gain score and teacher instructional practice scale score regressions. This appendix outlines the statistical models used in the analysis. In addition, it describes the imputation procedure used for the regression covariates and provides evidence supporting the use of gain scores as a dependent variable in achievement analyses.

In the models in which mean achievement gain scores are estimated as conditional upon a set of child-, teacher-, and school-level characteristics, the indeterminacy in separating teacher and school variance components is addressed by including only teacher effects in the models, which capture both the teacher- and school-level variability. This provides an implicit definition of a "teacher effect" as a teacher effect plus any school effect. This two-level hierarchical conditional means model can be represented as follows, using notation commonly employed to describe hierarchical models of this type (e.g., Raudenbush and Bryk 2002; Singer 1998):

$$Y_{ij} = \alpha_{0j} + \boldsymbol{x_{ij}}\,\gamma + \varepsilon_{ij}$$
$$\alpha_{0j} = \pi_{00} + \boldsymbol{z_j}\,\beta + \rho_{0j}$$

where $j = 1, \ldots, J$ indexes teachers (and schools taken together), $i = 1, \ldots, N$ indexes children, Y_{ij} is the individual student gain score in reading or mathematics, α_{0j} is the mean achievement gain of students of teacher j, $\boldsymbol{x_{ij}}$ is a vector of child-level variables, γ is the associated vector of coefficients, ε_{ij} is the random "child effect," π_{00} is the grand mean achievement gain, $\boldsymbol{z_j}$ is a vector of teacher-level variables, β is the associated vector of coefficients, and ρ_{0j} is the random "teacher and school effect." Substituting the second equation into the first yields the following equation that can be estimated:

$$Y_{ij} = \pi_{00} + \mathbf{z_j}\,\beta + \boldsymbol{x_{ij}}\,\gamma + \rho_{0j} + \varepsilon_{ij}$$

Statistical associations between Y and the teacher- and school-level variables contained in **z**, adjusted for the child-level variables contained in **x**, are reflected in β.

The teacher-level analysis employs a similar two-level hierarchical linear model to estimate the relationship of teacher scores on the specific instructional practice scales to their qualifications, taking into account the clustering of teachers within schools. A random intercept was included to pick up school-level effects. The two-level model described above applies, with the substitution of teachers for children and schools for teachers. The characteristics of teachers used as covariates in the analysis are years of kindergarten teaching experience, whether they hold a full teaching certification, whether they possess a master's degree, the number of courses completed in methods of teaching reading or mathematics, and whether they work full or part time.

All linear mixed models were estimated via restricted maximum likelihood using the HLM statistics package. Survey weights were applied to the level 1 variables, and multiple imputation

was used. The child-level analyses used the weight BYCW0, and the teacher-level analyses used the weight B2TW0.

Imputation Procedures

Missing data at the student, teacher, and school levels were imputed using the R package "norm" implementing the methods of Schafer (1997). A single imputed dataset was generated as follows. First, missing school-level data were imputed. Then, conditional on the complete school-level data, any missing teacher-level data were imputed. Finally, conditional on the complete teacher- and school-level data, the missing student-level data were imputed. This resulted in one complete dataset at all levels. This procedure was carried out five times, resulting in five replicate datasets. In all cases, convergence diagnostics of the imputation models were carefully checked to ensure that the missing data were being sampled from the correct distributions. Also, in all cases, imputed discrete variables were renormalized to legitimate values of such variables. All key analyses were carried out on all of the five imputed datasets. The resulting estimates, standard errors, and significance calculations were based on combining quantities via the appropriate multiple imputation procedures (Rubin 1987), which account for the uncertainty due to missing values.

The Use of Gain Scores

The dependent variables in the achievement analyses are gain scores, which represent the differences between the IRT fall and spring scale scores for the reading and mathematics assessments. Using gain scores as the dependent variable, rather than spring scores with fall scores as a covariate, allows results to be presented in terms of progress made during the year, regardless of where along the continuum that progress is made. There are long-standing concerns about the unreliability of gain scores in the measurement literature, although these concerns have more recently been shown to be unfounded and based on faulty assumptions (e.g., Gottman and Rushe 1993; Williams and Zimmerman 1996). Rogosa and Willet (1983) show that gain-score reliabilities are strong when individual differences between pretest and posttest are substantial, as is the case in most longitudinal assessment applications (including the fall and spring kindergarten ECLS-K assessments). Maris (1998) argues that regression toward the mean is not a legitimate argument against using gain scores; nor is pretest measurement error a concern, unless assignment into independent variable groups is determined from pretest performance (which is not the case in ECLS-K). Additionally, the use of IRT scale scores and the adaptive testing approach used in ECLS-K limit the concern that gain scores may be unreliable due to floor and ceiling effects (Rock and Pollack 2002).

Appendix D:
Sample Statistics for Variables Used in Analyses

Table D-1. Unweighted sample statistics for variables used in child-level analyses: School year 1998–99

Variable	Sample size	Mean	Standard deviation	Minimum	Maximum
Total Observations	16,308	†	†	†	†
Child-level variables					
Mathematics achievement gain score	16,284	8.13	5.05	-26.76	31.60
Reading achievement gain score	15,494	10.00	6.17	-13.80	43.63
Female	16,307	0.49	0.50	0	1
Race/ethnicity					
Asian/Pacific Islander	15,665	0.06	0.23	0	1
Black, non-Hispanic	15,665	0.14	0.35	0	1
Hispanic	15,665	0.17	0.38	0	1
Other, non-Hispanic race	15,665	0.04	0.20	0	1
Socioeconomic status	15,711	0.04	0.79	-5	3
Age at fall assessment	16,301	68.49	4.46	45.77	96.50
Days between tests	16,308	185.95	21.08	115.00	262.00
Repeating kindergarten in 1998–99	15,242	0.04	0.21	0	1
Primary home language non-English	15,717	0.12	0.32	0	1
School-level variables					
School locale					
Suburban	16,308	0.47	0.50	0	1
Rural	16,308	0.12	0.33	0	1
School control					
Private	16,308	0.22	0.42	0	1
Enrollment					
0-149 students	16,143	0.07	0.25	0	1
150-299 students	16,143	0.20	0.40	0	1
300-499 students	16,143	0.28	0.45	0	1
500-749 students	16,143	0.29	0.46	0	1
750 or more students	16,143	0.16	0.36	0	1
Percent minority enrollment	15,879	37.26	34.86	0	100
Percent free-lunch-eligible enrollment	10,458	29.24	28.20	0	100
Region					
West	16,308	0.23	0.42	0	1
Midwest	16,308	0.26	0.44	0	1
South	16,308	0.32	0.47	0	1

See notes at end of table.

Table D-1. Unweighted sample statistics for variables used in child-level analyses: School year 1998–99—Continued

Variable	Sample size	Mean	Standard deviation	Minimum	Maximum
Teacher-level variables					
90 or more minutes reading instruction per day	15,256	0.49	0.50	0	1
60 or more minutes mathematics instruction per day	15,162	0.55	0.50	0	1
Full-day kindergarten	15,218	0.56	0.50	0	1
Years of kindergarten teaching experience	16,209	9.09	7.71	1	38
Part-time employment status	14,342	0.06	0.24	0	1
Full certification	15,650	0.86	0.35	0	1
Coursework in methods of teaching reading					
1 course	15,257	0.15	0.36	0	1
2 courses	15,257	0.25	0.43	0	1
3 courses	15,257	0.19	0.40	0	1
4 courses	15,257	0.10	0.30	0	1
5 courses	15,257	0.05	0.23	0	1
6 or more courses	15,257	0.23	0.42	0	1
Coursework in methods of teaching mathematics					
1 course	15,215	0.25	0.43	0	1
2 courses	15,215	0.26	0.44	0	1
3 courses	15,215	0.19	0.39	0	1
4 courses	15,215	0.08	0.28	0	1
5 courses	15,215	0.03	0.17	0	1
6 or more courses	15,215	0.13	0.33	0	1

† Not applicable.
NOTE: Variables with a minimum of 0 and a maximum of 1 were coded as dummy variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Table D-2. Unweighted sample statistics for variables used in teacher-level analyses: School year 1998–99

Variable	Sample Size	Mean	Standard Deviation	Minimum	Maximum
Total Observations	3,305	†	†	†	†
Teacher-level variables					
Instructional practice scales					
Reading and writing skills	3,068	3.19	1.29	1	6
Phonics	3,084	5.53	0.50	1	6
Didactic instruction	3,078	3.10	1.31	1	6
Comprehension	3,061	4.69	0.96	1	6
Reading and writing activities	3,084	4.36	1.11	1	6
Student-centered reading instruction	3,082	3.42	0.85	1	6
Mixed-achievement reading grouping	3,066	3.98	1.42	1	6
Numbers and geometry	3,073	4.47	0.72	1	6
Advanced numbers and operations	3,057	3.48	1.35	1	6
Traditional practices and computation	3,075	3.19	0.99	1	6
Student-centered mathematics instruction	3,062	3.55	0.96	1	6
Mixed-achievement mathematics grouping	3,050	3.65	1.49	1	6
Measurement and advanced topics	3,073	2.72	0.84	1	6
Part-time employment status	2,769	0.06	0.23	0	1
Years of kindergarten teaching experience	3,280	8.20	7.46	0	38
Full certification	3,171	0.85	0.36	0	1
Master's degree	3,092	0.36	0.48	0	1
Coursework in methods of teaching reading					
1 course	3,098	0.16	0.37	0	1
2 courses	3,098	0.25	0.43	0	1
3 courses	3,098	0.19	0.39	0	1
4 courses	3,098	0.11	0.31	0	1
5 courses	3,098	0.05	0.22	0	1
6 or more courses	3,098	0.21	0.41	0	1
Coursework in methods of teaching mathematics					
1 course	3,081	0.27	0.44	0	1
2 courses	3,081	0.26	0.44	0	1
3 courses	3,081	0.18	0.38	0	1
4 courses	3,081	0.08	0.27	0	1
5 courses	3,081	0.03	0.18	0	1
6 or more courses	3,081	0.12	0.32	0	1
90 or more minutes reading instruction per day	2,983	0.54	0.50	0	1
60 or more minutes mathematics instruction per day	2,969	0.59	0.49	0	1
Full-day kindergarten	2,846	0.61	0.49	0	1
School-level variables					
School locale					
Suburban	3,305	0.49	0.50	0	1
Rural	3,305	0.09	0.29	0	1
School control					
Private	3,305	0.13	0.34	0	1
Enrollment					
0-149 students	3,286	0.05	0.23	0	1
150-299 students	3,286	0.11	0.32	0	1
300-499 students	3,286	0.24	0.43	0	1
500-749 students	3,286	0.32	0.47	0	1
750 or more students	3,286	0.27	0.45	0	1
Percent minority enrollment	3,242	44.65	35.83	0	100
Region					
West	3,305	0.27	0.44	0	1
Midwest	3,305	0.19	0.39	0	1
South	3,305	0.39	0.49	0	1

† Not applicable.

NOTE: Variables with a minimum of 0 and a maximum of 1 were coded as dummy variables.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Table D-3. Bivariate correlations of kindergarten teacher qualification variables: School year 1998-99

Teacher qualifications	Years of kindergarten teaching experience	Part-time employment status	Full certification	Master's degree	Coursework in methods of teaching reading	Coursework in methods of teaching mathematics
Years of kindergarten teaching experience	1.00	†	†	†	†	†
Part-time employment status	-.10	1.00	†	†	†	†
Full certification	.26	-.01	1.00	†	†	†
Master's degree	.21	-.02	.17	1.00	†	†
Coursework in methods of teaching reading	.20	.01	.21	.22	1.00	†
Coursework in methods of teaching mathematics	.25	.01	.22	.16	.75	1.00

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Appendix E: Standard Errors From Regression Models

Table E-1. Standard errors of the standardized regression coefficients for the regressions of kindergarten student achievement gains in reading on instructional practice scales: School year 1998–99

Characteristic	Instructional practice scales (standardized regression coefficients)							
	Analysis including phonics	Analysis including didactic instruction	Analysis including reading and writing skills	Analysis including reading and writing activities	Analysis including compre-hension	Analysis including mixed-achievement grouping	Analysis including student-centered instruction	Full analysis
Intercept	0.075	0.076	0.076	0.077	0.078	0.079	0.079	0.071
Teacher-level variables								
90 minutes or more reading instruction per day	0.025	0.025	0.025	0.026	0.025	0.025	0.025	0.025
Full-day kindergarten	0.029	0.029	0.029	0.029	0.030	0.029	0.030	0.028
Instructional practice scales								
Phonics	0.014	†	†	†	†	†	†	0.014
Didactic instruction	†	0.013	†	†	†	†	†	0.014
Reading and writing skills	†	†	0.014	†	†	†	†	0.017
Reading and writing activities	†	†	†	0.013	†	†	†	0.016
Comprehension	†	†	†	†	0.013	†	†	0.016
Mixed-achievement grouping	†	†	†	†	†	0.012	†	0.013
Student-centered instruction	†	†	†	†	†	†	0.014	0.016
School-level variables								
School locale[1]								
Suburban	0.027	0.027	0.027	0.028	0.028	0.028	0.028	0.027
Rural	0.041	0.041	0.041	0.042	0.042	0.042	0.042	0.041
School control[2]								
Private	0.043	0.042	0.042	0.045	0.043	0.044	0.043	0.042
Enrollment[3]								
150-299 students	0.067	0.066	0.067	0.069	0.069	0.070	0.070	0.060
300-499 students	0.066	0.065	0.066	0.068	0.069	0.069	0.069	0.060
500-749 students	0.066	0.066	0.066	0.068	0.069	0.069	0.069	0.060
750 or more students	0.069	0.068	0.069	0.071	0.071	0.072	0.072	0.062
Percent minority enrollment	0.059	0.059	0.059	0.059	0.060	0.059	0.060	0.057
Percent free-lunch-eligible enrollment	0.079	0.081	0.079	0.080	0.079	0.079	0.079	0.077
Region[4]								
West	0.039	0.038	0.038	0.039	0.039	0.039	0.039	0.037
Midwest	0.035	0.034	0.034	0.035	0.035	0.035	0.035	0.033
South	0.035	0.036	0.035	0.036	0.036	0.036	0.036	0.035

See notes at end of table.

Table E-1. Standard errors of the standardized regression coefficients for the regressions of kindergarten student achievement gains in reading on instructional practice scales: School year 1998–99—Continued

Characteristic	Instructional practice scales (standardized regression coefficients)							
	Analysis including phonics	Analysis including didactic instruction	Analysis including reading and writing skills	Analysis including reading and writing activities	Analysis including compre-hension	Analysis including mixed-achievement grouping	Analysis including student-centered instruction	Full analysis
Child-level variables								
Female	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016
Race/ethnicity[5]								
Asian/Pacific Islander	0.047	0.046	0.046	0.047	0.047	0.047	0.047	0.046
Black, non-Hispanic	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032
Hispanic	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033
Other, non-Hispanic race	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.043
Socioeconomic status	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010
Age at fall assessment	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
Days between tests	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
Kindergarten repeater	0.047	0.046	0.047	0.047	0.046	0.046	0.046	0.047
Non-English primary home language	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044

† Not applicable.
[1]The reference group for school locale is urban.
[2]The reference group for school control is public.
[3]The reference group for school size is less than 150 students.
[4]The reference group for school region is Northeast.
[5]The reference group for race/ethnicity is White, non-Hispanic.
NOTE: Regression coefficients are presented in table 1 of the main report. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Table E-2. Standard errors of the standardized regression coefficients for the regressions of kindergarten student achievement gains in reading on teacher qualifications: School year 1998–99

Characteristic	Teacher characteristics (standardized regression coefficients)
Intercept	0.118
Teacher-level variables	
90 minutes or more reading instruction per day	0.025
Full-day kindergarten	0.030
Years of kindergarten teaching experience	0.018
Experience squared	0.009
Part-time employment status	0.052
Full certification	0.043
Master's degree	0.030
Coursework in methods of teaching reading[1]	
1 course	0.102
2 courses	0.100
3 courses	0.099
4 courses	0.103
5 courses	0.114
6 or more courses	0.100
School-level variables	
School locale[2]	
Suburban	0.027
Rural	0.042
School control[3]	
Private	0.044
Enrollment[4]	
150-299 students	0.068
300-499 students	0.068
500-749 students	0.067
750 or more students	0.070
Percent minority enrollment	0.059
Percent free-lunch-eligible enrollment	0.080
Region[5]	
West	0.039
Midwest	0.035
South	0.036

See notes at end of table.

Table E-2. Standard errors of the standardized regression coefficients for the regressions of kindergarten student achievement gains in reading on teacher qualifications: School year 1998–99—Continued

Characteristic	Teacher characteristics (standardized regression coefficients)
Child-level variables	
Female	0.016
Race/ethnicity[6]	
Asian/Pacific Islander	0.047
Black, non-Hispanic	0.032
Hispanic	0.033
Other, non-Hispanic race	0.044
Socioeconomic status	0.010
Age at fall assessment	0.009
Days between tests	0.011
Kindergarten repeater	0.047
Non-English primary home language	0.044

[1]Courses are in methods of teaching reading; the reference group is zero courses.
[2]The reference group for school locale is urban.
[3]The reference group for school control is public.
[4]The reference group for school size is less than 150 students.
[5]The reference group for school region is Northeast.
[6]The reference group for race/ethnicity is White, non-Hispanic.
NOTE: Regression coefficients are presented in table 2 of the main report. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Table E-3. Standard errors of the standardized regression coefficients for the regressions of kindergarten teachers' reading instructional practice scales on selected teacher qualifications and school characteristics: School year 1998–99

Characteristic	Instructional practice scales (standardized regression coefficients)						
	Phonics	Didactic instruction	Compre-hension	Reading and writing activities	Reading and writing skills	Student-centered instruction	Mixed-achievement grouping
Intercept	0.522	0.238	-0.405	0.277	0.195	0.234	0.218
Teacher-level variables							
90 minutes or more reading instruction per day	0.054	0.047	0.054	0.047	0.048	0.045	0.052
Full-day kindergarten	0.127	0.074	0.074	0.082	0.069	0.065	0.072
Part-time employment status	0.225	0.120	0.101	0.118	0.098	0.115	0.135
Years of kindergarten teaching experience	0.039	0.033	0.036	0.031	0.029	0.036	0.034
Experience squared	0.023	0.018	0.021	0.018	0.016	0.020	0.019
Full certification	0.117	0.074	0.078	0.066	0.065	0.068	0.075
Master's degree	0.060	0.045	0.056	0.045	0.046	0.050	0.051
Coursework in methods of teaching reading[1]							
1 course	0.351	0.162	0.137	0.164	0.113	0.169	0.166
2 courses	0.337	0.164	0.130	0.164	0.112	0.171	0.156
3 courses	0.334	0.170	0.133	0.166	0.121	0.175	0.162
4 courses	0.337	0.167	0.139	0.165	0.120	0.169	0.169
5 courses	0.348	0.197	0.154	0.181	0.137	0.199	0.183
6 or more courses	0.336	0.169	0.144	0.168	0.119	0.181	0.164
School-level variables							
School locale[2]							
Suburban	0.141	0.076	0.065	0.082	0.062	0.067	0.064
Rural	0.166	0.112	0.113	0.117	0.101	0.101	0.097
School control[3]							
Private	0.118	0.109	0.114	0.100	0.098	0.099	0.105
Enrollment[4]							
150-299 students	0.216	0.141	0.144	0.140	0.117	0.130	0.133
300-499 students	0.215	0.135	0.151	0.137	0.121	0.119	0.129
500-749 students	0.225	0.138	0.150	0.140	0.118	0.120	0.132
750 or more students	0.230	0.140	0.159	0.143	0.127	0.129	0.137
Percent minority enrollment	0.148	0.156	0.143	0.116	0.110	0.107	0.138
Percent free-lunch-eligible enrollment	0.180	0.216	0.165	0.161	0.139	0.156	0.185
Region[5]							
West	0.144	0.100	0.103	0.113	0.092	0.088	0.091
Midwest	0.100	0.096	0.090	0.103	0.079	0.094	0.090
South	0.168	0.101	0.089	0.116	0.093	0.091	0.087

[1]Courses are in methods of teaching reading; the reference group is zero courses.
[2]The reference group for school locale is urban.
[3]The reference group for school control is public.
[4]The reference group for school size is less than 150 students.
[5]The reference group for school region is Northeast.
NOTE: Regression coefficients are presented in table 3 of the main report. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Table E-4. Standard errors of the standardized regression coefficients for the regressions of kindergarten student achievement gains in mathematics on instructional practice scales: School year 1998–99

Characteristic	Instructional practices scales (standardized regression coefficients)						
	Analysis including numbers and geometry	Analysis including advanced numbers and operations	Analysis including traditional practices and computation	Analysis including measure-ment and advanced topics	Analysis including mixed-achievement grouping	Analysis including student-centered instruction	Full analysis
Intercept	0.067	0.066	0.064	0.067	0.067	0.067	0.063
Teacher-level variables							
60 minutes or more mathematics instruction per day	0.025	0.025	0.024	0.025	0.024	0.024	0.025
Full-day kindergarten	0.028	0.028	0.028	0.029	0.028	0.028	0.028
Instructional practice scales							
Numbers and geometry	0.012	†	†	†	†	†	0.014
Advanced numbers and operations	†	0.011	†	†	†	†	0.012
Traditional practices and computation	†	†	0.012	†	†	†	0.013
Measurement and advanced topics	†	†	†	0.013	†	†	0.016
Mixed-achievement grouping	†	†	†	†	0.013	†	0.014
Student-centered instruction	†	†	†	†	†	0.012	0.015
School-level variables							
School locale[1]							
Suburban	0.024	0.024	0.024	0.024	0.024	0.024	0.024
Rural	0.038	0.038	0.038	0.038	0.039	0.038	0.038
School control[2]							
Private	0.036	0.036	0.035	0.036	0.036	0.036	0.036
Enrollment[3]							
150-299 students	0.057	0.056	0.053	0.057	0.057	0.057	0.052
300-499 students	0.055	0.054	0.051	0.055	0.055	0.055	0.051
500-749 students	0.056	0.055	0.052	0.056	0.056	0.056	0.052
750 or more students	0.059	0.058	0.055	0.059	0.059	0.059	0.055
Percent minority enrollment	0.058	0.057	0.057	0.058	0.058	0.057	0.057
Percent free-lunch-eligible enrollment	0.076	0.076	0.075	0.077	0.077	0.076	0.077
Region[4]							
West	0.034	0.033	0.033	0.034	0.033	0.033	0.033
Midwest	0.029	0.029	0.029	0.029	0.029	0.029	0.029
South	0.033	0.033	0.032	0.033	0.033	0.033	0.033

See notes at end of table.

Table E-4. Standard errors of the standardized regression coefficients for the regressions of kindergarten student achievement gains in mathematics on instructional practice scales: School year 1998–99—Continued

Characteristic	Instructional practices scales (standardized regression coefficients)						
	Analysis including numbers and geometry	Analysis including advanced numbers and operations	Analysis including traditional practices and computation	Analysis including measure-ment and advanced topics	Analysis including mixed-achievement grouping	Analysis including student-centered instruction	Full analysis
Child-level variables							
Female	0.017	0.017	0.017	0.017	0.017	0.017	0.017
Race/ethnicity[5]							
Asian/Pacific Islander	0.050	0.050	0.049	0.050	0.050	0.050	0.050
Black, non-Hispanic	0.031	0.031	0.030	0.031	0.031	0.031	0.030
Hispanic	0.031	0.031	0.031	0.032	0.032	0.031	0.031
Other, non-Hispanic race	0.046	0.046	0.046	0.046	0.046	0.046	0.046
Socioeconomic status	0.010	0.010	0.010	0.010	0.010	0.010	0.010
Age at fall assessment	0.009	0.009	0.009	0.009	0.009	0.009	0.009
Days between tests	0.010	0.010	0.010	0.010	0.010	0.010	0.010
Kindergarten repeater	0.046	0.046	0.046	0.046	0.046	0.046	0.046
Non-English primary home language	0.037	0.036	0.036	0.036	0.036	0.036	0.036

† Not applicable.
[1]The reference group for school locale is urban.
[2]The reference group for school control is public.
[3]The reference group for school size is less than 150 students.
[4]The reference group for school region is Northeast.
[5]The reference group for race/ethnicity is White, non-Hispanic.
NOTE: Regression coefficients are presented in table 4 of the main report. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-use Data File, fall 1998 and spring 1999.

Table E-5. Standard errors of the standardized regression coefficients for the regression of kindergarten student achievement gains in mathematics on teacher qualifications: School year 1998–99

Characteristic	Teacher characteristics (standardized regression coefficients)
Intercept	0.091
Teacher-level variables	
60 minutes or more mathematics enrollment per day	0.024
Full-day kindergarten	0.029
Years of kindergarten teaching experience	0.017
Experience squared	0.009
Part-time employment status	0.052
Full certification	0.038
Master's degree	0.006
Coursework in methods of teaching mathematics[1]	
1 course	0.061
2 courses	0.062
3 courses	0.064
4 courses	0.072
5 courses	0.104
6 or more courses	0.068
School-level variables	
School locale[2]	
Suburban	0.024
Rural	0.038
School control[3]	
Private	0.037
Enrollment[4]	
150-299 students	0.057
300-499 students	0.054
500-749 students	0.056
750 or more students	0.059
Percent minority enrollment	0.057
Percent free-lunch-eligible enrollment	0.075
Region[5]	
West	0.034
Midwest	0.029
South	0.034

See notes at end of table.

Table E-5. Standard errors of the standardized regression coefficients for the regression of kindergarten student achievement gains in mathematics on teacher qualifications: School year 1998–99—Continued

Characteristic	Teacher characteristics (standardized regression coefficients)
Child-level variables	
Female	0.017
Race/ethnicity[6]	
Asian/Pacific Islander	0.050
Black, non-Hispanic	0.031
Hispanic	0.032
Other, non-Hispanic race	0.046
Socioeconomic status	0.010
Age at fall assessment	0.009
Days between tests	0.010
Kindergarten repeater	0.047
Non-English primary home language	0.036

[1]Courses are in methods of teaching mathematics; the omitted category is zero courses.
[2]The reference group for school locale is urban.
[3]The reference group for school control is public.
[4]The reference group for school size is less than 150 students.
[5]The reference group for school region is Northeast.
[6]The reference group for race/ethnicity is White, non-Hispanic.
NOTE: Regression coefficients are presented in table 5 of the main report. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.

Table E–6. Standard errors of the standardized regression coefficients for the regressions of kindergarten teachers' mathematics instructional practice scales on selected teacher qualifications and school characteristics: School year 1998–99

Characteristic	Instructional practice scales (standardized regression coefficients)					
	Numbers and geometry	Advanced numbers and operations	Traditional practices and computation	Student-centered instruction	Mixed-achievement grouping	Measurement and advanced topics
Intercept	0.216	0.187	0.246	0.214	0.208	0.189
Teacher-level variables						
60 minutes or more mathematics instruction per day	0.051	0.046	0.047	0.049	0.055	0.045
Full-day kindergarten	0.073	0.066	0.075	0.072	0.074	0.066
Part-time employment status	0.128	0.105	0.129	0.123	0.128	0.108
Years of kindergarten teaching experience	0.035	0.032	0.029	0.034	0.034	0.031
Experience squared	0.020	0.018	0.017	0.019	0.019	0.017
Full certification	0.078	0.071	0.069	0.082	0.080	0.063
Master's degree	0.052	0.050	0.044	0.055	0.053	0.046
Coursework in methods of teaching mathematics[1]						
1 course	0.136	0.099	0.109	0.150	0.122	0.099
2 courses	0.132	0.102	0.108	0.141	0.121	0.101
3 courses	0.135	0.107	0.110	0.148	0.123	0.103
4 courses	0.137	0.117	0.119	0.153	0.136	0.111
5 courses	0.173	0.139	0.139	0.176	0.160	0.144
6 or more courses	0.143	0.109	0.118	0.160	0.133	0.117
School-level variables						
School locale[2]						
Suburban	0.067	0.064	0.085	0.066	0.068	0.069
Rural	0.098	0.114	0.115	0.092	0.100	0.099
School control[3]						
Private	0.104	0.103	0.115	0.094	0.110	0.099
Enrollment[4]						
150-299 students	0.137	0.127	0.157	0.122	0.138	0.133
300-499 students	0.135	0.129	0.149	0.122	0.137	0.133
500-749 students	0.132	0.133	0.153	0.123	0.141	0.136
750 or more students	0.143	0.142	0.158	0.128	0.145	0.143
Percent minority enrollment	0.118	0.117	0.146	0.098	0.138	0.127
Percent free-lunch-eligible enrollment	0.173	0.142	0.183	0.127	0.183	0.160
Region[5]						
West	0.100	0.096	0.114	0.086	0.095	0.100
Midwest	0.094	0.097	0.108	0.083	0.095	0.096
South	0.093	0.095	0.116	0.091	0.091	0.096

[1]Courses are in methods of teaching mathematics; the reference group is zero courses.
[2]The reference group for school locale is urban.
[3]The reference group for school control is public.
[4]The reference group for school size is less than 150 students.
[5]The reference group for school region is Northeast.
NOTE: Regression coefficients are presented in table 6 of the main report. The instructional practice scales, percent minority enrollment, percent free-lunch eligible students, socioeconomic status, age at fall assessment, and days between tests were analyzed as continuous variables.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Base Year Restricted-Use Data File, fall 1998 and spring 1999.