Teacher Effects and the Achievement Gap: Do Teacher and Teaching Quality Influence the Achievement Gap Between Black and White and High- and Low-SES Students in the Early Grades?

LAURA M. DESIMONE

University of Pennsylvania

DANIEL LONG

Wesleyan University

Background/Context: *Although there is relative agreement on the pattern of the achievement gap, attributing changes in the gap to schooling is less clear. Our study contributes to understanding potential teacher and teaching effects on achievement and inequality.*

Purpose/Objective/Research Question/Focus of Study: *We intend our work to contribute to understanding the school's role in addressing the achievement gap. We investigate the extent to which specific aspects of teacher quality (degree in math, experience, certification, math courses, and professional development) and teaching quality (time spent on math instruction and conceptual, basic procedural, and advanced procedural instruction) influence mathematics achievement growth and the achievement gap between White and Black students and low- and high-SES students in kindergarten and first grade.*

Research Design, Data Collection and Analysis: *In this secondary analysis, we examine the first four waves of data from the National Center for Education Statistics' Early Childhood Longitudinal Study (2000), a nationally representative longitudinal sample of students who were kindergartners in 1998. We use multilevel growth models to estimate relationships.*

Findings/Results: *We found evidence that lower achieving students are initially assigned to*

teachers who emphasize basic instruction, and higher achieving students are assigned teachers who emphasize more advanced instruction. The use of advanced procedural instruction and time spent on math were related to achievement growth for traditionally disadvantaged populations—Black students and low-SES students. Other types of instruction and teacher quality variables were not related to achievement growth.

Conclusions/Recommendations: *We found weak or no effects for teacher quality and type of instruction, which suggests that these aspects of teacher and teaching quality may operate as sorting variables. This may explain a part of the findings of past cross-sectional and gain studies that would likely interpret correlations between teachers and teaching as part of the effect of instruction. We found that low achievers tend to get teachers who spend less time on instruction, a variable we found significant in influencing achievement growth. If, as our study found, time on instruction matters, and disadvantaged students are more likely to get the weakest teachers who spend less time on instruction, we can identify an area in which schooling exacerbates the achievement gap but has the potential to ameliorate it.*

Inequality in education outcomes has been the target of research and policy at least since the War on Poverty in 1965. A central question has been whether schools mitigate or exacerbate the achievement gap between students of different racial/ethnic backgrounds and family income levels. Some argue that schools reproduce inequalities (Bourdieu & Passeron, 1977; Bowles & Gintis, 1976) through systematic differences within schools (e.g., tracking; Gamoran & Mare, 1989; Oakes, 1985) or between schools (e.g., teacher quality; Condron & Roscigno, 2003). Others believe that the schools act to decrease inequality (Cremin, 1951) by providing disadvantaged students with educational experiences they otherwise would not be exposed to at home.

Our study contributes to understanding the school's role in inequality by investigating the extent to which specific aspects of teacher and teaching quality influence student mathematics achievement growth and the achievement gap between White and Black students and low- and high-SES students in kindergarten and first grade, using a nationally representative sample of students, the Early Childhood Longitudinal Study (ECLS).

NATIONAL DATA AND TRENDS IN THE ACHIEVEMENT GAP

Recent analyses of national and other large-scale data have found clear patterns in the achievement gap among races/ethnicities. Reviews of the National Assessment of Education Progress (NAEP; Braswell et al., 2001) and other national data show that before formal schooling begins, Black children perform about half of a standard deviation lower than White children in mathematics, reading, and vocabulary. Furthermore, this gap

remains constant in reading and widens by an additional two-tenths of a standard deviation in mathematics and vocabulary by 12th grade (Phillips Crouse, & Ralph, 1998).

For mathematics, which we focus on in this study, Tate (1997) examined three national data sets—the NAEP, the National Education Longitudinal Study (NELS), and Scholastic Achievement Test (SAT) scores—and found that from 1973 to 1992, both White and Black students experienced positive growth in mathematics proficiency. However, there was still a large achievement gap in mathematics between the two groups, and White students outperformed Black students at each grade level. Hedges and Nowell (1998) found similar results by examining six large nationally representative surveys conducted between 1965 and 1992.

Studies examining the extent to which SES explains the race achievement gap are mixed. Lubienski (2002) analyzed NAEP data from 1990 to 2000 and found that Black–White gaps on math performance were significant at both the lowest and highest SES levels and that the lowest SES White students consistently scored equal to or higher than the highest SES Black students across three grades (fourth, eighth, and twelfth) in 1990 and 1996. In contrast, Fryer and Levitt (2004) analyzed the ECLS and modeled the gap such that SES variables explained much of the Black–White achievement gap in kindergarten and first grade.

LACK OF CLEAR EVIDENCE ABOUT THE SCHOOL'S ROLE IN THE ACHIEVEMENT GAP

Although there is relative agreement on the pattern of the achievement gap, attributing changes in the gap to schooling is less clear. One complication is that the gap between advantaged and disadvantaged students has been shown to increase during the summer recess (Alexander, Entwisle, & Olson, 2001; Cooper, Nye, Charlton, Lindsay & Greathouse, 1996; Heyns, 1978, 1987), which confounds the interpretation of the school's role when summer is not accounted for separately. Further, these same studies found that the pattern of gaps and growth across reading and mathematics can be quite different, as can results depending on how the gap is defined—for example, between White and Black and high- and low-SES students (Alexander et al., 2001). Further, out-of-school time varies substantially in the quality of educational experiences for the child (Hart & Risley, 1995), a variable that is difficult to include in models of the gap. Still another confounding factor is that regression toward the mean might explain the increasing of the gap—for example, Black students regress toward the Black student population mean, which is lower

than the student population mean for Whites (Porter, 2005).

The complexities of regression toward the mean, summer effects, different findings across subjects and for race and SES, and the quality and variation in out-of-school time make it difficult to determine the school's role in maintaining or improving the achievement gap (Phillips et al., 1998). The most recent empirical work that attempts directly to address the question of whether schools improve or magnify the achievement gap found that schools improve the achievement gap. In a careful analysis, Downey, von Hippel, and Broh (2004) used ECLS-K data, which is a nationally representative longitudinal sample of children who were kindergartners in 1998. They separated summer and school effects and showed that the Black–White achievement gap decreased during the school year. Additionally, Fryer and Levitt (2004) analyzed the kindergarten wave of the ECLS-K and concluded that a substantial part of the Black–White achievement gap might be due to Black students attending "worse" schools than White students.

Our study builds on this work by going inside the "black box" of schooling to examine the extent to which teachers and teaching influence the achievement gap in mathematics. The ECLS is ideal for bridging achievement gap trends and linking it with classroom instruction. In our study we link work on reform-oriented mathematics teaching (e.g., Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Cohen & Hill, 2001; Newman & Associates, 1996; Spillane & Zeuli, 1999) with larger scale work on the achievement gap (Downey et al., 2004; Fryer & Levitt, 2004; Phillips et al., 1998).

A focus on mathematics is justified given that U.S. students are achieving at alarmingly low levels in math compared to other countries (Porter, 2005; Schmidt et al., 2001; U.S. Department of Education, 2003), and teacher quality is a major contributor to the problem (Schmidt, McKnight, & Raizen, 1997). Focusing on a single subject, a common practice in studying teaching (e.g., Xue & Meisels, 2004), allows us greater control over the potentially confounding effects of a subject given consistent differences in the reading and mathematics achievement gaps across grades (Phillips et al., 1998). Further, teaching technologies and required competencies are different for different subjects, as are the issues involved in teacher quality, such as content knowledge and certification requirements.

DOES TEACHER OR TEACHING QUALITY MATTER?

In trying to improve overall achievement and decrease the achievement gap, a major focus of education reform efforts has been teacher and

teaching quality, an especially promising reform option (Porter, 2005). Several studies using value-added approaches to link teachers to student outcomes in elementary school suggest that the effects of teachers may be quite substantial (Rowan, Correnti, & Miller, 2002; Wright, Horn, & Sanders, 1997).

The ECLS provides us with an opportunity to examine the extent to which teacher characteristics and instruction are related to overall achievement and the achievement gap. The four main teacher and teaching quality features discussed in the literature, and which we focus our study on, are content knowledge, experience, certification and reform-oriented instruction. Below we highlight the major findings in these areas, as they pertain to links to student achievement.

TEACHER QUALITY AND STUDENT ACHIEVEMENT

Most of the research on teacher credentials published since the debut of the Coleman Report (Coleman et al., 1966) has focused on the link between teacher characteristics and student achievement. These studies found positive associations between student achievement and teacher knowledge, measured as their Scholastic Aptitude (or, more recently, Assessment) Test (SAT) or National Teacher Examination (NTE) score (Ballou, 1996; Ehrenberg & Brewer, 1994, 1995; Ferguson, 1991; Ferguson & Ladd, 1996; Mosteller & Moynihan, 1972; Strauss & Sawyer, 1986; Wright et al., 1997). Other studies found connections between student achievement and teacher knowledge proxies such as college major, number of courses, or amount of professional development taken in a subject area (Cohen & Hill, 2000; Darling-Hammond, 2000; Goldhaber & Brewer, 1997, 2000; Monk, 1994; Monk & King, 1994; Wenglinsky, 2000, 2002; Wiley & Yoon, 1995). Greenwald, Hedges, and Laine (1996) found, based on a meta-analysis, that teachers who attend better colleges and/or score higher on standardized tests produce greater gains in student achievement but are less likely to teach low-SES, Black, or Hispanic students. Teaching experience has been associated with achievement gains in high school mathematics (Fetler, 1999) and elementary mathematics (Murnane & Phillips, 1981; Rowan et al., 2002).

Studies of the relationship between teacher certification and student performance are more mixed (e.g., Darling-Hammond, Berry, & Thorenson, 2001; Goldhaber & Brewer, 2000). For example, Hawk, Coble, and Swanson (1985) found a positive relationship between mathematics achievement and teacher certification in secondary school, whereas Fetler (1999) found a negative correlation between mathematics and emergency credentials. Goldhaber and Brewer (2000) found no dif-

ference in mathematics achievement according to emergency or regular teacher certification for high school, and Rowan et al. (2002) found a similar lack of relationship at the elementary school level. These varying findings are likely in part because certification is operationalized quite differently across states.

INSTRUCTION AND STUDENT ACHIEVEMENT

Student opportunity to learn, defined as time spent on instruction in the classroom, has for several decades been shown to matter for student achievement (e.g., Carroll, 1963; Gamoran, Porter, Smithson, & White, 1997; Guarino, Hamilton, Lockwood, Rathbun, & Hausken, 2006). It is especially salient for disadvantaged students, who often do not receive high-quality education experiences outside of school (Alexander et al., 2001).

   In addition to the amount of time spent on academic content, in recent years, there has been a focus on the importance of the type of instruction that teachers use in mathematics (Cohen, McLaughlin, & Talbert, 1993; Elmore, Peterson, & McCarthy, 1996; Lampert, 1990; Milesi & Gamoran, 2005). Our choice about how to measure mathematics instruction is grounded in this literature and reflects current reform efforts in mathematics. The characteristics of effective teaching and teachers are varied and complex, and there is no firm consensus on what "good" teaching looks like (see Loveless, 2001). Thus, we developed our measures of teaching from current work in the teaching and learning of mathematics that seeks to increase the emphasis on conceptual learning goals (e.g., reasoning, estimation, conjecture; Cohen & Ball, 1990; National Commission on Teaching and America's Future, 1996; National Council of Teachers of Mathematics, 1989; Spillane & Zeuli, 1999) and decrease the emphasis on procedural learning goals (e.g., memorization, computation, routine problem-solving), which predominates instruction across American classrooms (Schmidt et al., 1997).

   Debates persist about the appropriate balance between conceptual and procedural instruction in mathematics; it has not yet been determined which mix of content with which students has what effect over what duration of time under what circumstances (see Gamoran, Secada, & Marrett, 2000; Loveless, 2001; Shouse, 2001). However, many studies have documented achievement benefits from increased use of conceptual techniques in mathematics, from using different definitions of conceptual instruction, to studying different grade levels (e.g., Carpenter et al., 1989; Cobb et al., 1991; Cohen & Hill, 2000; Gamoran et al., 1997; Hamilton et al., 2003; Hiebert et al., 1996, 1997; Lee, Smith, & Croninger, 1997; Silver

& Lane, 1995). Research also shows that conceptual techniques might be especially beneficial to disadvantaged students (e.g., Knapp, Shields, & Turnbull, 1992; Smith, Lee, & Newmann, 2001), but that compared with their high- and midachieving counterparts, low-achieving students receive less conceptual and more procedural instruction on average (Knapp & Shields, 1990; Kozma & Croninger 1992; Levine, 1988; Smith et al., 2005). Some studies, though, offer evidence in support of an emphasis on direct, procedural instruction (e.g., Geary, 2001; Slavin, Madden, Karweit, Livermon, & Dolan, 1990), especially in the early grades (D'Agostino, 2000).

A few studies have attempted to examine the effects of different types of instruction on the achievement gap. One study examined the differential impact of reform (inquiry-based) and traditional (teacher-centered) types of instruction on Black and Hispanic students' math performance by using a subsample of 190 Black and 174 Hispanic students from NELS: 88 (Manswell Butty, 2001). Results showed no significant differences between the two types of instruction on 10th-grade students' math achievement; however, for 12th-grade students, reform-oriented instruction was significantly more effective than traditional teaching. Another study examined the impact of reform-based instruction on closing fourth-grade students' achievement gap in mathematics by using the 2000 NAEP (Wenglinsky, 2004). In this study, Wenglinsky distinguished two types of achievement gaps: a within-school gap and a between-school gap. The results showed that when instructional practices were taken into account, the within-school gaps disappeared, but the between-school gaps remained unchanged. In addition, the study indicated that some instructional practices were beneficial to all students, irrespective of race. For example, increasing class time spent on math, emphasizing routine problems, and emphasizing geometry were found to be beneficial to all students. Some practices were found to have negative impacts on all student achievement. Those practices included frequent testing, working on projects, and emphasizing facts. The practices particularly beneficial to African American students differed somewhat from those beneficial to the whole student body. For Black students, the most beneficial practice was the emphasis on topics of measurement, and the most detrimental practice was taking tests.

CONTRIBUTIONS OF OUR STUDY

Our focus on *teacher* quality is based on the belief that certain teacher background characteristics are related to better teaching. The focus on *teaching* quality is based on the belief that better instruction leads to

improved student achievement. Drawing on achievement gap and teacher quality research, we focus our inquiry on contrasting procedural and conceptual approaches to teaching and several key teacher quality indicators—specifically, certification, whether the teacher is inexperienced (less than two years of teaching experience), content knowledge as proxied by degree in mathematics, mathematics courses, and professional development taken in mathematics. Figure 1 reflects our conceptual framework for the study.

Figure 1. Conceptual Framework for the Study

Context: Limited English proficiency, school
SES, and race/ethnicity
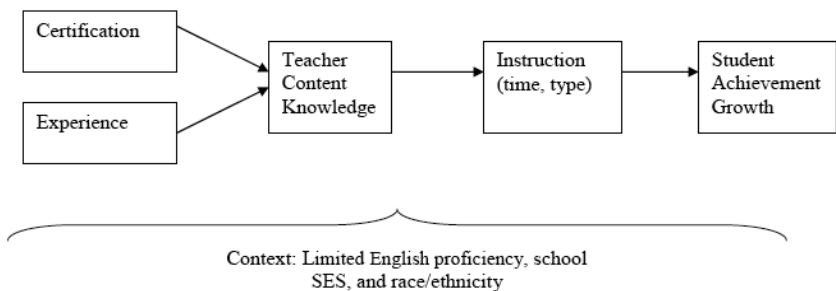
The ECLS provides an opportunity to build on previous work by examining the same students over time in order to identify the process through which teachers might affect achievement and the achievement gap. Our study has several strengths that build on and extend previous work: (1) we use nationally representative longitudinal data that allow growth modeling, (2) we use a hierarchical framework to account for the fact that students are nested within teachers, and time points are nested within students, (3) we take into account summer learning curves to make a stronger separation between school and out-of-school effects, much like Downey et al. (2004), and (4) we examine the initial teacher and teaching quality distribution as well as change over time to separate initial correlations due to teacher assignment from effects of teachers and teaching on growth. This last modification allows us to account for research that has shown that students from low-income homes are more likely to be taught by inexperienced teachers who are not certified and do not have a degree in the content area in which they are teaching (e.g., Goldhaber & Brewer, 2000; Ingersoll, 2002). Research also suggests that high-poverty students are more likely to have teachers who rely predominantly on basic/procedural rather than conceptual/higher order instruction (Barr, Wiratchai, & Dreeben, 1983; Desimone, Smith, & Frisvold, 2007; Gamoran, 1986; Smith, Desimone, & Ueno, 2005),

though high-poverty students do not necessarily have lower conceptual than procedural achievement (Desimone, Smith, Hayes, & Frisvold, 2005). We examine initial teacher assignments to see if they correspond to this pattern, and then look at whether those characteristics affect student cognitive growth. Finally, (5) we categorize mathematics instruction both as opportunity to learn (full-day kindergarten and time spent on mathematics) and into a three-part typology of types of instruction, which reflects current debates in the mathematics community about the efficacy of a relative emphasis on conceptual versus procedural approaches to teaching math. This allows us to test on a national sample whether particular types of instruction are more or less effective for advantaged and disadvantaged students' achievement growth.

Our study builds on a recent NCES gains analysis of ECLS data that linked teacher quality and student achievement in kindergarten (Guarino et al., 2006). We extend this work by measuring growth rather than just one-year gains, including first grade and accounting for summer effects, and using more comprehensive measures of teachers' learning experiences (including professional development) and credentials (differentiating among alternative, emergency, and high and low). We also focus on mathematics content rather than pedagogy, which has been shown to be more weakly related to achievement (Pellegrino, Baxter, & Glaser, 1999; Porter, Kirst, Osthoff, Smithson, & Schneider, 1993). Further, we look at effects of teacher quality and instruction on overall achievement as well as the achievement gap. Previous studies have used a multitude of race/ethnicity and income categories when examining the achievement gap. Here we focus on the most common racial gap, between Black and White students, and the most commonly examined income gap, between high- and low-SES students.

RESEARCH QUESTIONS

Our inquiry focuses on looking inside schooling to examine whether teacher and teaching quality matters for overall achievement and narrowing the achievement gap between White and Black and low- and high-SES students in kindergarten and first grade. We focus on three main questions. First we ask, *What is the distribution of teacher and teaching quality during the first year of kindergarten?* Here we hypothesize, based on previous research, that low-achieving students are more likely to be assigned to teachers with less experience, lower credentials, and less content knowledge and that teachers spend less time on instruction and are more likely to use procedural rather than conceptual approaches to instruction. Our second question asks, *To what extent do teacher quality, time spent on instruc-*

tion, and type of instruction predict growth in student achievement in kinder-
garten and first grade? Our third question analyzes the answer to the sec-
ond question in the context of inequality, asking, *To what extent do teacher
and teaching quality narrow the Black–White and low-/high-SES achievement
gap?*

DATA AND SAMPLE

We examine the first four waves of data from the National Center for
Education Statistics' (NCES, 2000) Early Childhood Longitudinal Study-
Kindergarten Cohort (ECLS-K), a nationally representative longitudinal
sample of students who were kindergartners in 1998. The kindergarten
sample is based on a national sample of schools with kFindergarten pro-
grams. Because the ECLS followed students, teachers and schools were
sampled in the first grade only if they included one or more ECLS-K chil-
dren in their classrooms (NCES, 2002b). This study examines kinder-
garten and first-grade teachers within schools using teacher and principal
surveys and student achievement scores from the "restricted use" version
of the ECLS. These data allow the linking of students to teachers and
schools. The ECLS provides data on a national multistage probability
sample of approximately 19,000 kindergartners and first graders in 3,000
classrooms in 1,000 schools. The Department of Education used a dual-
frame multistage sampling design in which 100 primary sampling units of
counties or groups of counties were selected. Private and public schools
were sampled separately within each of the chosen primary sampling
units, and about 23 kindergartners were selected from each of the sam-
pled schools. Students were followed from fall of kindergarten until fifth
grade, with a refresher sample in first grade. Specifically, the first wave of
the ECLS consisted of fall and spring kindergartner achievement tests,
teacher surveys (93% response rate), principal surveys (69%), and par-
ent interviews (85%). In fall 1999, a subsample of first graders was given
achievement tests. All students were tested again in spring 1999, 2001
(third grade), and 2003 (fifth grade). Teacher, administrator, and parent
data collections were also administered again in these years. Private
school kindergartners and Asian students were oversampled. This sam-
pling design allows comparisons among students by race and ethnicity
and by socioeconomic status (Tourangeau, Nord, Lê, Pollack, & Atkins-
Burnett, 2006).

   We analyze students who have taken at least one wave of the mathemat-
ics assessment test. This reduced our sample size from 21,399 to 19,730.
To conduct a growth curve analysis, we needed the date of entry into
kindergarten and the date of each of the four achievement tests. When

there was an assessment score without a date, we imputed the missing date from the median date of assessment for the student's school. If all the students in a given school were missing the dates of an assessment, we dropped observations at the assessment level. We dropped cases that had missing data on all four assessments. We only included cases in which we were able to match students to their teachers. These additional requirements further reduced our sample to 10,980 students in 2,164 schools. See Appendix A for how missing data affects our sample.

In this study, we used data from the kindergarten and first-grade waves and analyzed the full sample of students (N=10,980) as well as four subsamples: White students (n = 6,652); Black students (n = 1,447); low-SES students (n = 2,797), defined as students who are in the lowest SES quartile; and high-SES students (n = 2,736), defined as students who are in the highest SES quartile.

VARIABLES

We calculated teacher-level variables from teacher questionnaires and classroom averages of student characteristics, and we calculated school-level characteristics from administrator surveys and school averages of student characteristics.

ACHIEVEMENT SCORES

Our growth curve model is based on an analysis of student and classroom characteristics at three levels: multiple student assessment scores, nested within students, nested within schools. At the student assessment level, we examined the dependent variable of math achievement scaled using item response theory (IRT) so that it could be compared across all four tests (in the fall and spring of kindergarten and the fall and spring of first grade). Each test counts as a separate observation, such that a student who took all four assessment tests would have four different observations at level 1.

As we mentioned earlier, children's reading and math skills were tested on four occasions: spring and fall of kindergarten (1998–1999), and fall and spring of first grade (1999–2000). The first, second, and fourth tests were given to all available students; the third test in the fall of first grade

was given in a 30% random subsample of schools. Tests followed a two-stage format designed to reduce ceiling and floor effects. In the first stage, children took a brief "routing test" comprising items of a wide range of difficulty. In the second stage, children took a test containing questions of appropriate difficulty given the results of routing test (NCES, 2000). IRT was used to map children's answers onto a common 64-point scale for math. Few scores were clustered near the top or bottom of IRT scales, suggesting that ceiling and floor effects were minimized. In addition, the IRT scales improved validity and reliability by down-weighting questions with poor discrimination or high guessability (Rock & Pollack, 2002).

We paid careful attention to the dates on which tests were given, similar to Downey et al.'s (2004) study. We included variables for the number of days between the beginning of kindergarten and each test administration. So, the kindergarten (K), summer (S), and first grade (F) variables reflect how long the child was in each in grade (or the length of summer) before taking the assessment.

STUDENT VARIABLES

We included a series of control variables to measure SES and race at the student level. To measure SES, we used a measure calculated by NCES that consists of a weighted average of parent's education and income. This variable is standardized with a mean of 0 and a standard deviation of 1. We measured race with dummy variables for White, Black, Asian, Hispanic, Native American, and mixed race.[1] We also included controls for other factors that might be associated with achievement, such as the age of the student (in months) at the beginning of kindergarten and whether the student was in kindergarten for the first time in the 1998–1999 school year. Last, at the student level, we included days absent from kindergarten and first grade. The days-absent variable can also be viewed as a measure of opportunity to learn, given that absence from classes leads to less exposure to teaching and less of a chance to learn. See Table 1 for the descriptive statistics for all the level 1and level 2 covariates for the full sample and each of our subsamples (i.e., White, Black, and low and high SES).

Table 1. Descriptive Statistics

	Full Sample		White Sample		Black Sample		High SES Sample		Low SES Sample	
Level 1	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Math Achievement	31.29	12.57	33.03	12.53	27.21	11.43	35.72	12.35	26.88	12.01
Months in Kindergarten	213.26	95.1	212.21	94.69	214.79	94.52	211.93	95.09	214.48	95.69
Months on Summer Break	31.95	40.16	32.25	40.37	31.63	38.49	32.41	40.69	31.51	40.79
Months in First Grade	82.22	114.82	81.84	114.25	82.25	114.78	81.29	113.93	83.27	116
Sample Size	35110		21313		4691		8836		8877	
Level 2	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
"Socio Economic Status										
(SES), standardized"	0.05	0.77	0.22	0.73	-0.35	0.74	1.09	0.44	-0.81	0.45
White	0.61	0.49	1	0	0	0	0.76	0.43	0.38	0.49
Black	0.13	0.34	0	0	1	0	0.05	0.22	0.23	0.42
Hispanic	0.14	0.35	0	0	0	0	0.07	0.25	0.27	0.44
Asian	0.06	0.24	0	0	0	0	0.07	0.26	0.05	0.21
Native American	0.03	0.18	0	0	0	0	0.01	0.12	0.05	0.21
Mixed Race	0.02	0.15	0	0	0	0	0.03	0.17	0.02	0.14
1st time K	0.85	0.36	0.87	0.34	0.82	0.39	0.9	0.3	0.84	0.37
Age began Kindergarten	66.26	4.34	66.57	4.32	66.06	4.29	66.17	4.2	66.29	4.48
Days Absent in Kindergarten	9.16	12.3	8.08	9.1	9.71	10.34	7.2	8.86	11.82	16.3
Days Absent in First Grade	7.82	8.68	7.53	8.43	8.32	11.23	6.81	6.18	9.38	11.37
Sample Size	10980		6652		1447		2736			
Level 3										
Kindergarten										
Teacher Quality	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Experience (standardized)	0.04	0.96	0.07	0.96	0.02	0.95	0.07	0.96	0.04	0.98
New Teacher	0.1	0.31	0.1	0.3	0.11	0.31	0.1	0.3	0.12	0.32
Less than BA	0.01	0.07	0.01	0.08	0	0.07	0.01	0.1	0	0.06
BA	0.24	0.43	0.25	0.43	0.26	0.44	0.24	0.43	0.24	0.43
More than BA	0.59	0.49	0.58	0.49	0.55	0.49	0.59	0.49	0.6	0.49
High certification	0.57	0.49	0.58	0.49	0.54	0.5	0.58	0.49	0.57	0.49
Regular certification										
No certification	0.01	0.12	0.02	0.13	0.01	0.12	0.03	0.16	0.01	0.11
Emergency certification	0.07	0.25	0.06	0.24	0.07	0.25	0.06	0.24	0.07	0.26
Alternative certification	0.01	0.12	0.01	0.08	0.01	0.09	0.01	0.11	0.01	0.12
Professional development (?=.55)	0.01	0.61	0	0.61	0.02	0.61	0.01	0.61	0.01	0.62
Number of math courses	0	0.96	0.02	0.97	-0.03	0.92	0.01	0.98	-0.01	0.94
Kindergarten Instruction										
Full-day K	0.6	0.46	0.58	0.47	0.69	0.44	0.57	0.47	0.61	0.46
Minutes of math per day										
(10 min. increments)	18.12	10.68	17.47	10.67	18.6	10.85	17.22	10.21	18.71	10.7
Basic Procedural										
(in days per month)	9.74	4.78	9.34	4.76	10.06	4.99	9.45	4.65	9.87	4.68
Conceptual										
(in days per month)	5.78	4.33	5.66	4.36	6.28	4.6	5.79	4.32	5.65	4.22
Advanced Procedural										
(in days per month)	4.57	3.13	4.54	3.14	4.79	3.3	4.6	3.09	4.45	3.05
Kindergarten Controls	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
% private	0.1	0.3	0.12	0.33	0.07	0.26	0.19	0.39	0.06	0.23
% free lunch	30.68	24.37	25.22	20.27	39.4	27.36	22.45	19.26	37.08	25.76
Class > 27	0.13	0.33	0.11	0.31	0.13	0.33	0.13	0.33	0.14	0.34
10% or less LEP	0.74	0.42	0.8	0.38	0.78	0.4	0.81	0.37	0.7	0.44
% min. in class	0.38	0.34	0.25	0.24	0.52	0.33	0.28	0.28	0.45	0.36
% w/ Data on										
Kindergarten Teachers	0.93	0.25	0.93	0.26	0.92	0.27	0.94	0.23	0.94	0.25

Table 1. Descriptive Statistics (continued)

	Full Sample		White Sample		Black Sample		High SES Sample		Low SES Sample	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
First Grade Teacher Quality										
Experience	0.04	0.81	0.09	0.8	-0.01	0.83	0.11	0.82	0.02	0.8
New Teacher	0.09	0.24	0.09	0.23	0.12	0.27	0.08	0.22	0.1	0.25
Less than BA	0	0.06	0	0.06	0.01	0.07	0.01	0.08	0	0.03
BA	0.26	0.38	0.27	0.38	0.31	0.4	0.27	0.39	0.27	0.39
More than BA	0.58	0.44	0.61	0.43	0.54	0.44	0.62	0.43	0.58	0.44
High certification	0.07	0.22	0.08	0.23	0.1	0.26	0.08	0.22	0.07	0.22
Regular certification										
No certification	0.02	0.12	0.02	0.12	0.02	0.12	0.03	0.15	0.01	0.09
Emergency certification	0.06	0.21	0.06	0.19	0.07	0.22	0.06	0.2	0.07	0.22
Alternative certification	0.01	0.1	0.01	0.07	0.01	0.09	0.01	0.08	0.02	0.12
Professional development (?=.49)	-0.01	0.49	-0.02	0.48	0.06	0.52	-0.02	0.49	-0.01	0.5
Number of math courses	0	0.78	0	0.79	0	0.81	0.01	0.8	0.01	0.79
First Grade Instruction										
Minutes of math per day										
(10 min. increments)	23.39	11.62	23.9	10.58	24.16	11.61	23.79	10.01	23.93	11.56
Basic Procedural	8.56	4.79	8.65	4.4	8.88	4.81	8.55	4.22	8.73	4.8
Conceptual	7.57	4.51	7.79	4.25	7.94	4.58	7.84	4.07	7.5	4.49
Advanced Procedural	7.73	4.45	7.78	4.09	7.97	4.44	7.86	4	7.83	4.45
First Grade Controls										
% private	0.1	0.3	0.12	0.33	0.08	0.26	0.19	0.39	0.06	0.23
% free lunch	24.36	23.09	22.06	19.64	29.82	25.04	18.83	17.63	29.2	25.06
Class > 27	0.05	0.2	0.04	0.18	0.05	0.2	0.05	0.21	0.04	0.18
10% or less LEP	0.71	0.43	0.81	0.36	0.78	0.39	0.83	0.35	0.68	0.44
% min. in class	0.35	0.35	0.23	0.24	0.48	0.34	0.26	0.28	0.41	0.37
% w/ Data on										
First Grade Teachers	0.87	0.34	0.9	0.29	0.88	0.32	0.92	0.27	0.88	0.33
Sample Size	2164		1574		611		1048		1295	

TEACHER AND INSTRUCTION VARIABLES

At the teacher level, we examined three sets of variables for both kindergarten and first grade: teacher quality, instruction, and controls. We measured teacher characteristics with the following variables: years of experience teaching (standardized); a dummy variable for first- or second-year teachers to identify new teachers; teacher's education (whether they have a BA or graduate degree in mathematics); the level of teacher certification (high [permanent or long term certification], regular, emergency, alternative, or no certification); the number of college-level courses taken in mathematics; and the teacher's participation in professional development. We measured professional development with a composite consisting of an average of the following four standardized variables: (1) the teacher attended peer feedback meetings, (2) the

teacher attended or observed other schools, (3) the teacher attended classes and meetings to learn new skills, and (4) the teacher attended professional development workshops during the current academic year. The Cronbach's alpha for this professional development construct was .55 for kindergarten and .49 for first grade.

We measured exposure to instruction in several ways. For kindergartners, we used a dummy variable indicating full-day versus half-day kindergarten. For both kindergarten and first grade, we used a measure of the number of minutes the teacher spent on math instruction per day (in 10-minute increments) and three measures of the days per month the teacher spent on basic procedural teaching, conceptual teaching, and advanced procedural teaching.

To develop the three types of instruction, we used previous research (as reviewed earlier) to sort the ECLS items into categories of instruction. We then tested our conceptions using factor analysis and tested the final composites using Cronbach's alpha.

The *basic procedural* teaching measure consisted of the average of how many days per month the students were taught the following math skills: telling time, adding single-digit numbers, subtracting single-digit numbers, writing all numbers from 1 to 10, recognizing and naming geometric shapes, and making, copying, or extending patterns. We also included a question about the following math activity: How many days per month did the children in your class count out loud? The Cronbach's alpha for these variables was .74 for kindergarten teachers and .69 for first-grade teachers.

Conceptual teaching consisted of the average of the number of days per month that students were taught the following skills: estimating quantities, estimating probabilities, and writing math equations to solve word problems. We also included in this measure the following question about math activities: How many days per month did the children in your class work on math problems that reflect real life situations? The Cronbach's alpha for conceptual teaching was .65 for both kindergarten and first grade.

The *advanced procedural* teaching composite measured the extent to which the teacher taught the following skills: place value, reading two-digit numbers, mixed operations, recognizing fractions, recognizing the value of coins and currency, counting by 2s, 5s, and 10s, counting beyond 100, writing all the numbers between 1 and 100, reading three-digit numbers, adding two-digit numbers, carrying numbers in addition, and subtracting two-digit numbers. The Cronbach's alpha for the advanced procedural teaching construct was .75 for kindergarten and .82 for first grade.

SCHOOL- AND CLASS VARIABLES

Last, we included controls indicating whether a student was in private school, the percent of students in the school who qualified for free lunch, the percentage of minorities in the classroom, a dummy variable for classes with more than 27 students, a dummy variable for classes with less than 10% limited-English-proficiency (LEP)[2] students, and a dummy variable for whether we had data for the kindergarten or first-grade teacher. (See Table 1 for the descriptive statistics for all the school and classroom variables.) Variables describing conditions at the class level were assigned to the corresponding teachers and thus were categorized as teacher-level variables in our models.

We tested for collinearity at the teacher level and found that there were no notable correlations between the teacher-level covariates for either first grade or kindergarten (see Appendix B).

METHODS

We estimated knowledge and learning rates using a multilevel growth model (Raudenbush & Bryk, 2002; Singer & Willett, 2003). In this model, we viewed tests (level 1) as nested within children, and children (level 2) as nested within schools (level 3).

This model allowed us to examine the *distribution of teacher and teaching quality during the first year of* k*indergarten* and to estimate the effects of *teacher quality, time spent on instruction, and type of instruction on the growth in student achievement in kindergarten and first grade.*

GROWTH MODEL

To estimate the growth curve, we used a three-level multilevel model in hierarchical linear modeling (HLM; see equations below). Level 1 consists of the IRT math score as the dependent variable. We modeled the growth in math achievement with three covariates of the days since the beginning of kindergarten, designed to separate academic-year growth from summer gains or losses. The kindergarten slope (K) consists of the days since the beginning of kindergarten for assessments between the beginning and end of kindergarten and 9 months later. The summer slope (S) consists of assessment dates between 9 months and 12 months since the beginning of kindergarten. The first-grade (F) slope consists of assessment dates 12 months after the beginning of kindergarten.

Our model extrapolates the scores that would have been obtained on the last day of kindergarten and the first day of first grade. There might

be a slight bias if learning speeds up or slows down at the beginning and end of the school year, though Downey et al. (2004) found that learning rates are approximately constant for much of the school year. Our model makes these extrapolations by using information about the date of each test relative to the first and last days of school.

Level 2 consists of student family background (SES), student race (RACE), student family background (SES), whether this was the first time that a student entered kindergarten (First Time in K.) and student age for the intercept. The slope for kindergarten includes race, family background, whether the student was in a full-day kindergarten (Full day K.), and the number of days absent in kindergarten. The summer slope consists of race and family background. The slope for first grade includes race, family background , and days absent in first grade.

Level 3 consists of a series of teacher quality, instruction, and control variables for both kindergarten and first grade. The kindergarten covariates at level 3 influence both the level 1 intercept and the growth of kindergarten achievement. The first-grade covariates are modeled on the first-grade slope.

Level 1: $Y = \pi_0 + \pi_1 (K) + \pi_2 (S) + \pi_3 (F) + e$

Level 2: $\pi_0 = B_{00} + B_{01} (SES) + B_{02} (RACE) + B_{03} (First\ Time\ in\ K.)$
$+ B_{04} (AGE) + r_0$

$\pi_1 = B_{10} + B_{11} (SES) + B_{12} (RACE) + B_{13} (Days\ Absent\ in\ K.)$
$+ B_{14} (Full\ day\ K.)$

$\pi_2 = B_{20} + B_{21} (SES) + B_{22} (RACE)$

$\pi_3 = B_{30} + B_{31} (SES) + B_{32} (RACE) + B_{33} (Days\ Absent\ in$
First Grade)

Level 3: $B_{00} = {}_{001} + {}_{002}(K.\ Teacher\ Quality) + {}_{003}(K.\ Instruction) +$
${}_{004}(K.\ Controls) + u_{00}$

$B_{01} = {}_{01}$

$B_{02} = {}_{02}$

$B_{03} = {}_{03}$

$B_{04} = {}_{04}$

$B_0 = {}_{00} + {}_{00} (K.\ Teacher\ Quality) + {}_{00} (K.\ Instruction) + {}_{00}$
(K. Controls)

$B_{10} = {}_{11}$

$B_{12} = {}_{12}$

$B_{13} = {}_{13}$

$B_{20} = {}_{00}$

$B_{21} = {}_{21}$

$B_{22} = {}_{21}$

$$B_{30} = {}_{00}\ {}_{00}\ \text{(F. Teacher Quality)} + {}_{00}\ \text{(F. Instruction)} + {}_{00}\ \text{(F. Controls)}$$

$$B_{31} = {}_{10}$$
$$B_{32} = {}_{20}$$
$$B_{33} = {}_{30}$$

The equations above jointly model the sorting of students, teachers, and classroom achievement on initial achievement and the influences on three learning rates: kindergarten, summer, and first grade. We estimated the equations using HLM for the full model and for the White, Black, high-SES, and low-SES subsamples (see Tables 2 and 3).

The intercept in this growth model (π_0 and B_{00}) can be interpreted as the initial sorting of students of different abilities into schools and classrooms. In other words, the covariates ${}_{001}$, ${}_{002}$, and ${}_{003}$ measured the selection effects of unmeasured characteristics that matched student achievement with different types of kindergarten teaching and different types of kindergarten teachers. For example, the positive and significant effect of the private school covariate on the intercept could be interpreted to indicate that more higher achieving students begin kindergarten in private schools than in public schools. In contrast, the slopes for kindergarten (π_1 and B_{10}) and first grade (π_3 and B_{30}) gave estimates of the effects of teaching and teacher characteristics on achievement growth. For example, the private school covariate for the slope is either negative or zero, meaning that private schools are not associated with an increase in achievement growth above the initial sorting of high-ability students into private schools.

This growth curve model has several advantages over ordinary regressions. In ordinary regressions, the estimated correlation between initial status and subsequent change is attenuated and negatively biased because measurement error is confounded with true variation in initial status and change (Blomqvist, 1977; Thomson, 1924). Our multilevel growth model avoids this bias by separating school and student-level variation from variation due to test-level measurement error.

MISSING VALUES

ECLS-K has a fair number of missing values. We assumed that values were missing at random (Allison, 2001; Little & Rubin, 2002). We dropped cases with missing data on the date of assessment and where we were not able to impute an assessment date from other students in the same school. However, randomly missing test scores are not problematic because our longitudinal models did not require that all children be

tested on all occasions (Raudenbush & Bryk, 2002; Singer & Willett, 2003). As a result, we were able to keep cases that had at least one out of the four assessments. For the covariates at levels 2 and 3, we used mean substitution for missing data and added a dummy variable (see Appendix A for the descriptive statistics for all the missing data dummies).[3]

Randomly missing predictors can produce bias and inefficiency. We addressed this potential problem by creating dummy variables for missing data and doing mean substitution for missing values. This produced unbiased estimates but slightly inefficient standard errors. We could have used a multiple imputation strategy (Allison, 2001; Rubin, 1987), which in nonnested data has the advantage of producing unbiased and efficient estimates. However, we tested a series of smaller multilevel models of achievement with the ECLS data and found almost identical coefficients and standard errors with both mean imputation and multiple imputation. Although multiple imputation has advantages for nonnested data, there is a great deal of uncertainty about how best to conduct multiple imputations with nested data. It is possible that imputation strategies can change the covariance matrices between the different levels of imputed data. The direction and magnitude of this bias is unknown. Because of this uncertainty, we used mean imputation instead.

Our analysis of the longitudinal data from the ECLS has several advantages over existing cross-sectional studies. First, our multiple measures of achievement at the beginning and end of kindergarten allow us to estimate both the initial distribution of achievement and achievement growth in order to distinguish between the effects of initial student characteristics and the effects of schooling on achievement growth. In contrast, cross-sectional studies suffer from an inability to distinguish the effects of initial student traits and schooling effects. Second, the multiple measures of achievement allow us to parse out the confounding summer learning effects. Third, the extensive covariates at both the kindergarten and first-grade levels allow us to account for the changing school and teacher traits that influence student achievement.

RESULTS

DIFFERENCES IN TEACHER QUALITY AND INSTRUCTION BY RACE AND SES (TABLE 1)

The descriptive statistics for the White versus Black sample and the high- versus low-SES sample show similarities in both teacher and teaching quality. The percentage of new teachers for all four subsamples is close to 10% in kindergarten and first grade. The number of teachers

with emergency credentials is near 6%, and the number with alternative credentials is near 1% for all four subgroups in kindergarten and first grade. The percentage of teachers with high (or advanced) certification is nearly identical for all subgroups in both kindergarten and first grade. In addition, levels of professional development for all subgroups are almost identical to the grand mean for professional development. There also are no statistically significant differences among the number of math courses, teacher education, and levels of teacher experience across subgroups. Contrary to findings from other studies, it appears that for kindergartners and first graders in the ECLS, Black, White, well-off, and poor students all have teachers with similar levels of experience, education, certification, professional development, and math coursework.

Though average levels of time spent on mathematics and emphasis on the three different types of instruction differed by subgroup, none of these differences was statistically significant. Black students' teachers reported spending slightly more time on math and on each of the three types of instruction compared with teachers of White students in both kindergarten and first grade. Low-SES students have slightly higher levels of minutes of instruction and days of basic instruction than high-SES students. In contrast, high-SES students have slightly higher levels of conceptual and algorithmic instruction. However, as we mentioned, none of these differences was statistically significant.

The large differences in subgroups are related to gaps in socioeconomic status between Black and White students and differential racial composition between high- and low-SES groups. In addition, the contextual controls of private schools, percent free lunch, and percent minority in the class are statistically different for all groups. There are also statistically significant differences in absenteeism for the subgroups.

Although there might not be much difference in the mean levels of instruction and teaching characteristics by subgroups, it is possible that there are notable differences in the effects of these characteristics in both their initial distribution across differential ability levels and on the growth in achievement. Next, we discuss the results concerning these two issues.

INITIAL STUDENT ASSIGNMENT (TABLES 2 AND 3)

The main findings we want to highlight in Tables 2 and 3 are that being new to the teaching profession and all three of our teaching measures were associated with the initial distribution of students across classrooms. Lower achieving kindergartners are more likely to have new teachers (b = -.64*), but none of the other teacher quality variables (certification, BA

or higher in mathematics, professional development participation) is significantly related to initial achievement levels, except that teachers who have taken more coursework in mathematics have students with lower achievement (b = -.17*). We suspect that taking coursework is a proxy for lack of prior content knowledge, so this negative relationship is consistent with the idea that weaker teachers get assigned to weaker students.

Teachers who more often use advanced procedural instruction (algorithms) (b = .08**) and conceptual approaches to mathematics (b = .08**) are more likely to have higher achieving kindergartners in their fall class. Teachers who favor more basic procedural approaches are more likely to have lower achieving kindergartners (b = -.06*). The subgroup analyses of Black, White, and high- and low-SES students show that the initial allocation of students to teachers generally followed the same patterns (e.g., conceptual and advanced procedural instruction were associated with higher initial achievement, and basic procedural instruction was associated with initially lower achieving students), but for the most part, coefficients were not significant, probably due in part to sample size limitations. One exception was that for high-SES students (Table 3, Model 4), having a teacher who spent more time on math instruction was associated with initially higher levels of achievement (b = 1.19**).

ACHIEVEMENT GROWTH AND SUMMER EFFECTS (TABLES 2 AND 3)

Moving to our growth models, we ask how teacher and teaching quality affects kindergarten and first-grade achievement growth. We separate summer growth from academic year growth to better isolate cognitive growth that occurs while the student is in school. Tables 2 and 3 show the extent to which teacher quality and instruction are related to student achievement growth in both kindergarten and first grade.

TEACHER QUALITY

In the full sample, reported in Table 2, several teacher quality variables are related to first-grade achievement growth, but none is associated with kindergarten growth. Specifically, first-grade achievement growth occurs at a *slower* rate if students have a teacher with less than a bachelor's degree (b = -.70**) and at a *faster* rate if the teacher has high certification (b = .30**) or alternative certification (b = .60*).

Subgroup analyses reveal similar but not identical patterns.[4] The Whites-only sample is similar to the full sample—for first graders, having a teacher with less than a BA in mathematics slows growth (b = -.87*), and having a teacher with high certification accelerates growth (b = .32*).

There are no teacher quality effects for kindergartners. For the Black-only sample, no teacher quality variables are significantly related to growth. For low-SES students, having a teacher with no certification is associated with an increase in cognitive growth in mathematics in kindergarten. For high-SES students in first grade, having a new teacher (b = -.53*) is associated with the slowing of achievement growth (see Models 2 and 3 in Table 2, and Models 4 and 5 in Table 3).

Table 2. Growth Curve Models for the Full Sample and the White and Black Samples

Initial Status	Model 1- Full Sample Coefficient	S.E.	P-value	Model 2 - Whites Coefficient	S.E.	P-value	Model 3 - Blacks Coefficient	S.E.	P-value
Intercept	-8.64	*(1.23)*	0.00	-9.73	*(1.71)*	0.00	-6.95	*(2.99)*	0.02
Teacher Quality									
Experience	-0.10	*(.10)*	0.28	-0.13	*(.13)*	0.33	0.06	*(.17)*	0.72
New Teacher	-0.64	*(.29)*	0.03	-0.74	*(.43)*	0.09	0.32	*(.56)*	0.57
Less than BA	0.06	*(1.00)*	0.96	-0.71	*(1.44)*	0.62	-0.10	*(1.40)*	0.94
More than BA	0.10	*(.20)*	0.60	0.02	*(.26)*	0.95	0.33	*(.32)*	0.30
High certification	-0.10	*(.21)*	0.62	-0.33	*(.28)*	0.25	0.15	*(.33)*	0.66
No certification	-0.39	*(.76)*	0.61	-0.60	*(1.25)*	0.63	-0.42	*(1.01)*	0.68
Emergency certification	-0.21	*(.33)*	0.53	-0.30	*(.45)*	0.50	-0.93	*(.55)*	0.09
Alternative certification	0.78	*(1.02)*	0.44	2.44	*(1.72)*	0.16	4.15	*(2.16)*	0.05
Professional development	0.18	*(.13)*	0.17	0.11	*(.17)*	0.52	0.04	*(.23)*	0.86
Number of math courses	-0.17	*(.08)*	0.04	-0.18	*(.11)*	0.09	-0.43	*(.16)*	0.01
Instruction									
Full-day K	0.02	*(.19)*	0.91	0.18	*(.25)*	0.48	-0.60	*(.38)*	0.11
Minutes of math	0.00	*(.01)*	0.81	0.00	*(.01)*	0.88	0.02	*(.02)*	0.21
Basic Procedural	-0.06	*(.03)*	0.03	-0.06	*(.04)*	0.12	0.01	*(.05)*	0.83
Advanced Procedural	0.08	*(.03)*	0.01	0.11	*(.04)*	0.02	0.07	*(.05)*	0.18
Conceptual	0.08	*(.02)*	0.00	0.07	*(.03)*	0.03	0.09	*(.04)*	0.02
Controls									
% private	1.77	*(.27)*	0.00	1.54	*(.32)*	0.00	2.96	*(.59)*	0.00
% free lunch	-0.03	*(.00)*	0.00	-0.04	*(.01)*	0.00	-0.01	*(.01)*	0.25
Class > 27	0.25	*(.31)*	0.43	-0.45	*(.44)*	0.30	0.72	*(.58)*	0.22
10% or less LEP	0.42	*(.25)*	0.09	-0.37	*(.42)*	0.37	-0.08	*(.62)*	0.90
% min. in class	-0.19	*(.34)*	0.58	0.24	*(.60)*	0.68	-0.42	*(.56)*	0.46
1st time K	2.06	*(.27)*	0.00	2.55	*(.41)*	0.00	0.38	*(.63)*	0.55
Black	-1.42	*(.22)*	0.00						
Hispanic	-1.85	*(.23)*	0.00						
Asian	0.66	*(.37)*	0.07						
Native American	-2.47	*(.44)*	0.00						
SES (sescont)	2.64	*(.11)*	0.00	2.88	*(.15)*	0.00	1.80	*(.22)*	0.00

(Table 2 continued on the next page)

Table 2. Growth Curve Models for the Full Sample and the White and Black Samples (continued)

Kindergarten Slope	Model 1- Full Sample			Model 2 - Whites			Model 3 - Blacks		
Intercept	0.05	(.00)	0.00	0.05	(.00)	0.00	0.05	(.00)	0.00
Teacher Quality									
Experience	0.05	(.00)	0.16	0.05	(.00)	0.27	-0.05	(.00)	0.59
New Teacher	0.09	(.00)	0.45	0.07	(.00)	0.65	-0.40	(.00)	0.12
Less than BA	-0.20	(.00)	0.60	-0.05	(.00)	0.89	-0.20	(.01)	0.81
More than BA	-0.04	(.00)	0.62	0.01	(.00)	0.90	-0.20	(.00)	0.36
High certification	0.02	(.00)	0.79	0.09	(.00)	0.37	0.03	(.00)	0.87
No certification	0.40	(.00)	0.12	0.40	(.00)	0.21	0.10	(.01)	0.82
Emergency certification	0.10	(.00)	0.43	0.20	(.00)	0.30	-0.20	(.00)	0.48
Alternative certification	-0.40	(.00)	0.23	-0.72	(.00)	0.09	-1.10	(.01)	0.09
Professional development	-0.09	(.00)	0.08	-0.09	(.00)	0.20	-0.10	(.00)	0.16
Number of math courses	-0.01	(.00)	0.79	0.01	(.00)	0.85	-0.03	(.00)	0.70
Instruction									
Full-day K (Khours2)	0.10	(.00)	0.13	0.19	(.00)	0.04	0.02	(.00)	0.91
Minutes of math	0.01	(.00)	0.10	0.003	(.00)	0.52	-0.001	(.00)	0.86
Basic Procedural	0.01	(.00)	0.13	0.01	(.00)	0.47	-0.01	(.00)	0.61
Advanced Procedural	0.03	(.00)	0.01	0.04	(.00)	0.01	0.04	(.00)	0.13
Conceptual	-0.01	(.00)	0.19	-0.01	(.00)	0.25	-0.01	(.00)	0.67
Controls									
% private	-0.20	(.00)	0.05	-0.19	(.00)	0.10	-0.20	(.00)	0.36
% free lunch	-0.002	(.00)	0.24	-0.004	(.00)	0.14	-0.01	(.00)	0.01
Class > 27	-0.08	(.00)	0.47	-0.15	(.00)	0.25	-0.46	(.00)	0.08
10% or less LEP	-0.06	(.00)	0.52	-0.06	(.00)	0.69	0.04	(.00)	0.88
% min. in class	-0.14	(.00)	0.29	0.008	(.00)	0.97	0.50	(.00)	0.13
Black	-0.50	(.00)	0.00						
Hispanic	-0.20	(.00)	0.04						
Asian	0.07	(.00)	0.62						
Native American	0.20	(.00)	0.33						
SES (sescont)	0.20	(.00)	0.00	0.17	(.00)	0.01	0.10	(.00)	0.31
Absent	-0.02	(.00)	0.00	-0.02	(.00)	0.00	-0.03	(.00)	0.00
Summer Slope									
Intercept	2.20	(.00)	0.00	2.30	(.00)	0.00	2.40	(.00)	0.00
SES cont	0.20	(.00)	0.20	0.10	(.00)	0.52	0.01	(.00)	0.98
MSEScount	0.30	(.01)	0.67	-0.07	(.01)	0.95	0.20	(.01)	0.86
Black	0.20	(.00)	0.55						
Hispanic	0.40	(.00)	0.37						
Asian	0.80	(.01)	0.11						
Native American	-0.70	(.01)	0.23						

Table 2. Growth Curve Models for the Full Sample and the White and Black Samples (continued)

1st Grade Slope	Model 1- Full Sample			Model 2 - Whites			Model 3 - Blacks		
Intercept	0.05	*(.00)*	0.00	0.04	*(.00)*	0.00	0.04	*(.00)*	0.00
Teacher Quality									
Experience	-0.01	*(.00)*	0.90	0.03	*(.00)*	0.56	-0.18	*(.00)*	0.06
New Teacher	-0.20	*(.00)*	0.10	-0.20	*(.00)*	0.33	-0.08	*(.00)*	0.79
Less than BA	-0.70	*(.00)*	0.01	-0.87	*(.00)*	0.04	-1.00	*(.01)*	0.25
More than BA	-0.10	*(.00)*	0.18	-0.05	*(.00)*	0.59	0.05	*(.00)*	0.79
High certification	0.30	*(.00)*	0.01	0.32	*(.00)*	0.02	0.40	*(.00)*	0.25
No certification	0.30	*(.00)*	0.14	0.30	*(.00)*	0.34	0.40	*(.01)*	0.56
Emergency certification	0.10	*(.00)*	0.51	0.10	*(.00)*	0.51	0.20	*(.00)*	0.46
Alternative certification	0.60	*(.00)*	0.04	0.30	*(.00)*	0.31	0.80	*(.01)*	0.34
Professional development	-0.02	*(.00)*	0.70	-0.08	*(.00)*	0.28	0.20	*(.00)*	0.13
Number of math courses	-0.02	*(.00)*	0.56	0.004	*(.00)*	0.94	0.01	*(.00)*	0.88
Instruction									
Minutes of math	0.07	*(.00)*	0.02	0.00	*(.00)*	0.45	00.02	*(.00)*	0.02
Basic Procedural	-0.01	*(.00)*	0.45	0.01	*(.00)*	0.72	-0.01	*(.00)*	0.68
Advanced Procedural	0.03	*(.00)*	0.04	0.02	*(.00)*	0.30	-0.02	*(.00)*	0.55
Conceptual	0.01	*(.00)*	0.23	0.01	*(.00)*	0.52	0.01	*(.00)*	0.63
Controls									
% private	-0.25	*(.00)*	0.01	-0.20	*(.00)*	0.15	-0.40	*(.00)*	0.12
% free lunch	-0.002	*(.00)*	0.15	0.00	*(.00)*	0.71	-0.01	*(.00)*	0.10
Class > 27	-0.20	*(.00)*	0.15	-0.09	*(.00)*	0.60	0.30	*(.00)*	0.32
10% or less LEP	-0.01	*(.00)*	0.92	0.10	*(.00)*	0.42	-0.07	*(.00)*	0.81
% min. in class	-0.10	*(.00)*	0.36	-0.10	*(.00)*	0.46	-0.65	*(.00)*	0.01
Black	-0.34	*(.00)*	0.01						
Hispanic	0.07	*(.00)*	0.60						
Asian	-0.59	*(.00)*	0.00						
Native American	-0.20	*(.00)*	0.30						
SES (sescont)	-0.16	*(.00)*	0.00	-0.15	*(.00)*	0.03	-0.08	*(.00)*	0.55
Absent	-0.01	*(.00)*	0.04	0.01	*(.00)*	0.03	0.002	*(.00)*	0.88
Variance Component									
Level 1 Variance	16.20			16.25			14.53		
Level 2 Variance	34.04			35.89			31.21		
df	8807.00			5074.00			832.00		
Chi-Square	53579.34			35243.73			10113.88		
Level 3 Variance	3.58			3.17			0.73		
df	2133.00			1543.00			580.00		
Chi-Square	3282.04			183.51			699.31		
Deviance	220475.06			134142.90			28882.86		
df	131			111.00			111.00		

Note: coefficients for kindergarten, summer and first-grade slopes were multiplied by 100

Table 3. Growth Curve Models for the Highest SES Quartile and the Lowest SES Quartile

Initial Status	Model 4 - High SES			Model 5 - Low SES		
	Coefficient	S.E.	P-value	Coefficient	S.E.	P-value
Intercept	-12.27	(2.74)	0.00	-5.68	(2.21)	0.01
Teacher Quality						
Experience	-0.22	(.23)	0.34	-0.05	(.14)	0.74
New Teacher	-1.41	(.72)	0.05	-0.20	(.45)	0.66
Less than BA	2.00	(1.93)	0.30	0.77	(1.88)	0.68
More than BA	-0.14	(.45)	0.75	0.40	(.31)	0.20
High certification	-0.15	(.51)	0.77	-0.46	(.33)	0.16
No certification	-1.62	(1.31)	0.22	-0.09	(.88)	0.92
Emergency certification	-0.94	(.87)	0.28	0.04	(.54)	0.94
Alternative certification	0.23	(1.51)	0.88	-0.72	(.99)	0.47
Professional development	0.31	(.30)	0.31	-0.04	(.20)	0.84
Number of math courses	-0.26	(.18)	0.16	0.03	(.13)	0.80
Instruction						
Full-day K	-0.02	(.02)	0.39	-0.41	(.31)	0.19
Minutes of math	1.19	(.44)	0.01	0.01	(.01)	0.37
Basic	-0.11	(.06)	0.07	-0.05	(.04)	0.25
Algorithm	0.15	(.07)	0.03	0.04	(.05)	0.38
Conceptual	0.10	(.05)	0.07	0.05	(.03)	0.13
Controls						
% private	1.30	(.45)	0.00	4.25	(.82)	0.00
% free lunch	-0.04	(.01)	0.00	-0.01	(.01)	0.04
Class > 27	-0.02	(.69)	0.98	0.66	(.48)	0.17
10% or less LEP	-0.07	(.62)	0.92	0.91	(.40)	0.03
% min. in class	-0.31	(.86)	0.72	-0.12	(.50)	0.81
1st time K	2.66	(.84)	0.00	1.42	(.43)	0.00
Black	-3.34	(.64)	0.00	-1.05	(.38)	0.01
Hispanic	-1.60	(.64)	0.01	-1.58	(.39)	0.00
Asian	1.19	(.67)	0.07	0.80	(.90)	0.38
Native American	-0.07	(1.30)	0.61	-3.57	(.64)	0.00
SES	2.13	(.38)	0.00	1.00	(.25)	0.00

(Table 3 continued on the next page)

Table 3. Growth Curve Models for the Highest SES Quartile and the Lowest SES Quartile (continued)

Kindergarten Slope	Model 4 - High SES			Model 5 - Low SES		
	Coefficient	S.E.	P-value	Coefficient	S.E.	P-value
Intercept	0.05	(.01)	0.00	0.05	(.00)	0.00
Teacher Quality						
Experience	0.10	(.00)	0.17	0.09	(.00)	0.10
New Teacher	0.41	(.00)	0.06	0.06	(.00)	0.76
Less than BA	-0.30	(.00)	0.34	0.09	(.01)	0.93
More than BA	0.03	(.00)	0.83	-0.10	(.00)	0.35
High certification	-0.02	(.00)	0.89	-0.03	(.00)	0.84
No certification	0.61	(.00)	0.06	1.10	(.00)	0.01
Emergency certification	0.06	(.00)	0.83	0.04	(.00)	0.87
Alternative certification	-0.72	(.00)	0.04	0.40	(.01)	0.44
Professional development	-0.10	(.00)	0.11	0.04	(.00)	0.65
Number of math courses	-0.01	(.00)	0.93	-0.02	(.00)	0.79
Instruction						
Full-day K	-0.10	(.00)	0.38	0.20	(.00)	0.13
Minutes of math	0.01	(.00)	0.30	0.01	(.00)	0.30
Basic Procedural	0.04	(.00)	0.03	0.02	(.00)	0.34
Advanced Procedural	0.01	(.00)	0.49	0.04	(.00)	0.08
Conceptual	-0.04	(.00)	0.05	-0.01	(.00)	0.37
Controls						
% private	-0.30	(.00)	0.03	-0.62	(.00)	0.01
% free lunch	0.003	(.00)	0.36	0.00	(.00)	0.05
Class > 27	-0.30	(.00)	0.12	-0.10	(.00)	0.50
10% or less LEP	0.30	(.00)	0.14	-0.55	(.00)	0.00
% min. in class	0.20	(.00)	0.52	-0.38	(.00)	0.10
Black	-0.65	(.00)	0.02	-0.52	(.00)	0.00
Hispanic	-0.48	(.00)	0.08	-0.47	(.00)	0.01
Asian	0.04	(.00)	0.87	0.30	(.00)	0.42
Native American	-0.60	(.01)	0.16	0.40	(.00)	0.19
SES (sescont)	0.08	(.00)	0.60	0.02	(.00)	0.89
Absent	-0.002	(.00)	0.64	-0.02	(.00)	0.00
Summer Slope						
Intercept	1.20	(.01)	0.12	2.00	(.01)	0.00
SES cont	0.80	(.01)	0.17	0.40	(.01)	0.36
MSEScount						
Black	1.00	(.01)	0.27	0.40	(.01)	0.52
Hispanic	1.80	(.01)	0.10	0.20	(.01)	0.68
Asian	-0.09	(.01)	0.92	0.70	(.01)	0.61
Native American	0.20	(.02)	0.92	-0.60	(.01)	0.50

Table 3. Growth Curve Models for the Highest SES Quartile and the Lowest SES Quartile (continued)

1st Grade Slope	Model 4 - High SES			Model 5 - Low SES		
	Coefficient	S.E.	P-value	Coefficient	S.E.	P-value
Intercept	0.05	(.00)	0.00	0.04	(.00)	0.00
Teacher Quality						
Experience	-0.10	(.00)	0.11	0.10	(.00)	0.13
New Teacher	-0.53	(.00)	0.02	-0.10	(.00)	0.68
Less than BA	0.05	(.00)	0.90	1.60	(.01)	0.06
More than BA	-0.003	(.00)	0.99	-0.10	(.00)	0.47
High certification	0.78	(.00)	0.00	0.20	(.00)	0.59
No certification	0.30	(.00)	0.32	-0.20	(.01)	0.79
Emergency certification	0.30	(.00)	0.21	-0.03	(.00)	0.89
Alternative certification	0.30	(.00)	0.50	0.86	(.01)	0.06
Professional development	-0.03	(.00)	0.77	-0.09	(.00)	0.42
Number of math courses	-0.04	(.00)	0.52	-0.06	(.00)	0.38
Instruction						
Minutes of math	0.003	(.00)	0.65	0.02	(.00)	0.01
Basic Procedural	-0.004	(.00)	0.84	-0.01	(.00)	0.63
Advanced Procedural	0.0006	(.00)	0.98	0.02	(.00)	0.49
Conceptual	0.02	(.00)	0.20	0.01	(.00)	0.45
Controls						
% private	-0.20	(.00)	0.22	-0.30	(.00)	0.26
% free lunch	0.001	(.00)	0.77	0.00	(.00)	0.10
Class > 27	-0.20	(.00)	0.19	-0.30	(.00)	0.29
10% or less LEP	-0.62	(.00)	0.00	0.20	(.00)	0.38
% min. in class	-0.40	(.00)	0.13	0.20	(.00)	0.43
Black	-0.40	(.00)	0.14	-0.41	(.00)	0.07
Hispanic	-0.20	(.00)	0.62	0.30	(.00)	0.14
Asian	-0.40	(.00)	0.13	-0.40	(.00)	0.34
Native American	-0.40	(.01)	0.45	-0.50	(.00)	0.13
SES (sescont)	-0.42	(.00)	0.02	-0.20	(.00)	0.17
Absent	-0.01	(.00)	0.14	-0.01	(.00)	0.23
Variance Component						
Level 1 Variance	16.68			16.80		
Level 2 Variance	36.62			32.52		
df	1680.00			1494.00		
Chi-Square	14622.66			12968.97		
Level 3 Variance	3.17			3.49		
df	1017.00			1264.00		
Chi-Square	1244.18			1631.78		
Deviance	55808.75			55926.40		
df	127.00			127.00		

Note: coefficients for kindergarten, summer and first-grade slopes were multiplied by 100

TEACHING QUALITY

As we explained earlier, we used two categories of instruction: (1) time spent on instruction, proxied by whether a student attended full-day kindergarten and minutes spent on mathematics instruction, and (2) type of instruction, measured by a typology reflecting the research on alternative approaches to mathematics—basic or lower level procedural, advanced procedural, and conceptual.

Time spent on instruction. Full-day kindergarten was not associated with achievement growth in any of the models. Minutes spent on mathematics instruction was associated with achievement growth for first graders in the full sample (b = .07*), the Black-only sample (b = .02*), and the low-SES sample (b=.02**).

Type of instruction. Advanced procedural instruction was associated with achievement growth in kindergarten in the full sample (b = .03**) and the White-only sample (b = .04**) (it was marginally significant at the p < .10 level for the low-SES sample, b =.04+). For first graders, advanced procedural instruction was associated with mathematics achievement growth for the full sample (b = .03*) only. Basic procedural and conceptual instruction were only significantly related to achievement growth for kindergartners in the high-SES sample. Here, basic instruction was associated with an acceleration of achievement growth (b = .04*), whereas conceptual instruction was associated with a decrease in achievement growth (b = -0.04*). Still, the overall trend of high-SES students having faster achievement growth still holds.

MAGNITUDE OF EFFECTS

Does teacher or teaching quality help close the achievement gap? How much would it take to make a substantial contribution to narrowing the gap in the early grades? Here we translate one of our key findings—that time spent on instruction is significantly related to achievement growth for Black students and students from low-SES families—into the context of the achievement gap.

Figure 2 shows the achievement gap at five key time points measured on the ECLS: the start of kindergarten, spring of kindergarten, the summer before first grade, fall of first grade, and spring of first grade. The Black–White math achievement gap, controlling for teacher characteristics, instruction, and classroom differences, is 1.42 points at the beginning of kindergarten and increases by 104% to 2.9 at the end of kindergarten. The gap at the beginning of first grade is 2.7 points, and by the end of first grade, the gap increases by 35%, to 3.64 points. To put

this in perspective, the overall achievement gain from the beginning of kindergarten until the end of kindergarten is 27.5 points for all racial/ethnic groups. White students experience a gain of 27.8, which is 101% of the total gain in achievement versus Black students, who gain 25.7, 93% of the average gain in achievement between the beginning of kindergarten and the end of first grade.

Figure 2. Growth Curve of Math Achievement by Race and Ethnicity
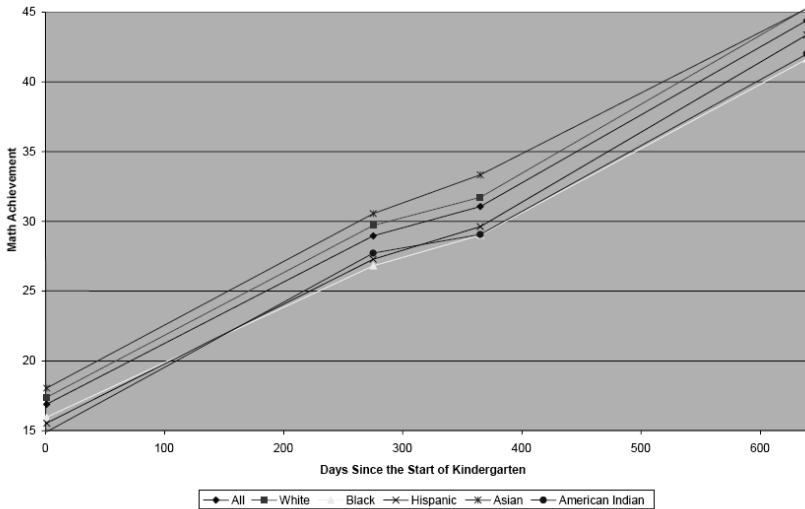
Based on these findings about minutes of instruction and the predicted growth in achievement from increasing minutes of math instruction (see Appendix C), an increase in 100 minutes of math instruction for Black students per day during kindergarten and first grade would decrease the Black–White gap by 10% by the end of first grade. An increase in 3 hours of instruction would decrease the gap by about 20%. This means that an additional 3 hours of instruction a day for 2 years would increase achievement for Black students from 25.7 points over 2 years to 26.4 points of math achievement growth, a 3% increase in achievement. This means that an increase of 29.3 points of math achievement, or a 14% increase in achievement over kindergarten and first grade, is needed to close the Black–White achievement gap.

Compared with increasing the number of minutes spent on mathematics instruction, additional advanced math instruction would lead to larger gains in achievement (see Appendix C, Panel D). An increase of 4 days per month of advanced math instruction for Black students during kindergarten and first grade would, by the end of first grade, decrease

the Black–White gap by 20%. An increase in 10 days of advanced instruction per month would decrease the gap by about 55%. This means that an additional four days of advanced instruction per month for 2 years would increase achievement from 25.7 points over 2 years to 27.2 points of math achievement growth, a 6% increase in achievement. This is still less than the 29.3 points in math achievement growth needed to close the Black–White achievement gap.

These predictions assume that the effects of instruction are the same for Black and White students and examine what would be the impact of a treatment if the means of measured teacher characteristics, types of instruction, SES, and classroom traits were the same for Black and White students. However, in making policy recommendations, it is worth examining what the effects of instruction would be if the means and effects varied by subgroup. Table 2, Model 2 shows the results of the growth curve analysis of the Black subsample. In Appendix D and Figure 3, we examine the predicted growth curve for Black students if both the means and effects are allowed to vary. This leads to an increase of 24.6 points in achievement from the beginning of kindergarten until the end of first grade. An addition of 10 days of advanced instruction per month would lead to an increase in achievement of 25.2 points, a 2% increase in achievement, and an increase of 3 hours of math instruction would increase overall achievement to 25.6, or an increase of 4%.

Figure 3. The Effect of Advanced Teaching and Minutes of Math Instruction on the Growth Curve in Math Achievement for Blacks vs. Whites
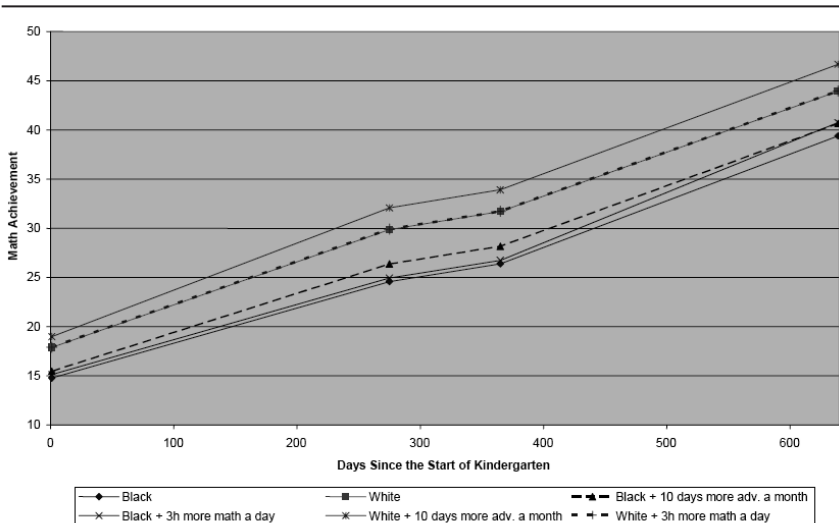
Figure 4 shows the SES achievement gap at the same five key time points as Figure 2: the start of kindergarten, the spring of kindergarten, the summer before first grade, the fall of first grade, and the spring of first grade for different SES percentiles. The gap between the top 75th SES percentile and the bottom 25th SES percentile, controlling for teacher characteristics, instructional differences, and classroom differences, is 2.61 points at the beginning of kindergarten and increases by 18% to 3.08 at the end of kindergarten. The gap at the beginning of first grade is 3.27 points, but by the end of first grade, the gap decreases by 14%, to 2.83 points. To put this in perspective, the overall achievement gain from the beginning of kindergarten until the end of kindergarten is 27.5 points for all SES groups. The top 75th percentile have a gain of 27.6, which is 100.3% of the total gain in achievement for the full sample versus the bottom 25th percentile, which gained 27.4 points, or 99.6% of the average gain in achievement between the beginning of kindergarten and the end of first grade for the full sample. It is worth noting that in contrast to the Black–White gap, the major differences in the high- and low-SES gap occur before entering kindergarten and remain much more constant during kindergarten and first grade.

**Figure 4. Growth Curve of Math Achievement by SES Percentile**
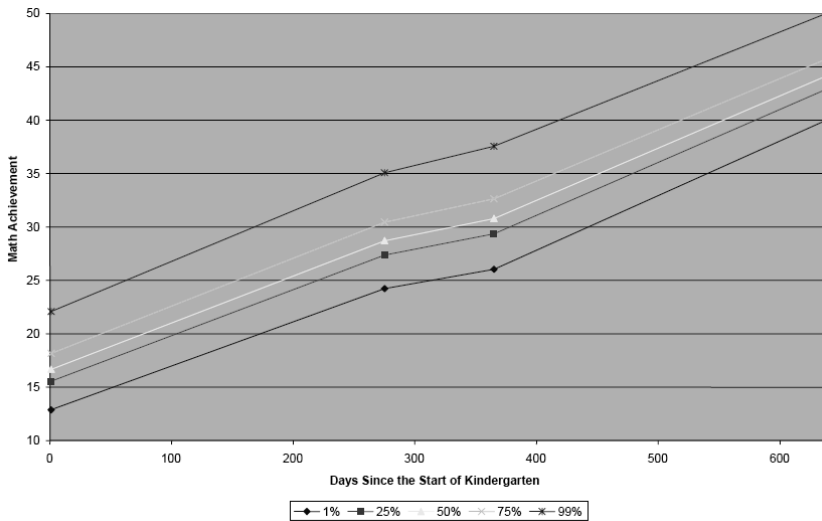
Based on these findings about minutes of instruction and the predicted growth in achievement from increasing minutes of math instruction (see Appendix C, Panel C), an increase in 100 minutes of math instruction

per day for students in the 25th percentile during kindergarten and first grade would, by the end of first grade, decrease the 25th/75th percentile SES gap by 13%. An increase in 3 hours of instruction would decrease the gap by about 23%. This means that an additional 3 hours of instruction a day for 2 years would increase achievement from 27.4 points over 2 years to 28.1 points of math achievement growth, a 2% increase in achievement.

Based on these findings, additional advanced math instruction would lead to larger gains in achievement than an increase in minutes of math instruction per day (see Appendix C, Panel D). An increase of 4 days of advanced math instruction per month for students in the 25th percentile during kindergarten and first grade would, by the end of first grade, decrease the SES gap by 22%. An increase in 10 days of advanced instruction per month would decrease the SES gap by about 55%. This means that an additional 10 days of advanced instruction per month for 2 years would increase achievement from 27.4 points over 2 years to 28.2 points of math achievement growth, a 3% increase in achievement.

These predictions assume that the effects of instruction are the same for high- and low-SES groups, and they examine treatment impact if the means of teacher characteristics, types of instruction, SES, and classroom traits are the same for all SES groups. Table 3, Model 5 shows the results of the growth curve analysis of the lowest SES quartile. In Appendix D and Figure 5, we examine the predicted growth curve for the lowest SES

Figure 5. The Effect of Advanced Teaching and Minutes of Math Instruction on the Growth Curve in Math Achievement for the Highest and Lowest SES quartiles
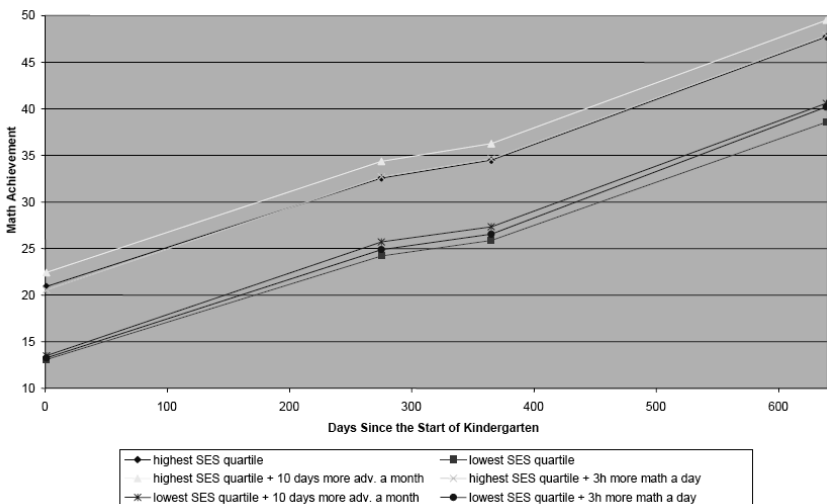
quartile if both the means and effects are allowed to vary. This leads to an increase in 25.4 points in achievement from the beginning of kindergarten until the end of first grade. An addition of 10 days of advanced instruction per month would lead to an increase in achievement of 27.1 points, a 6% increase in achievement, and an increase in 3 hours of math instruction would increase overall achievement to 26.94, or an increase in 5%.

DISCUSSION

We designed our study to investigate hypotheses about the extent to which key teacher and teaching quality variables influence overall achievement growth and the achievement gap in the early grades. In working toward these goals, our study has both strengths and weaknesses that should be considered in the interpretation of results.

FACTORS TO CONSIDER IN INTERPRETATION

Because neither teachers nor students are randomly assigned to classrooms, attributing teacher and teaching quality to student achievement gains is not straightforward. We account for some of this by examining initial allocation patterns and controlling for them and then measuring growth. Still, unobserved (omitted) variables are always a problem in nonexperimental designs, and results should be interpreted accordingly.

Our measures of teacher quality and instruction rely on teacher self-reports, which are often called into question. A careful look at the literature here, however, shows that survey measures of teaching, especially composite measures like the ones that we used in this study, are effective in describing and distinguishing among different types of teaching practices (Mayer, 1999). These survey measures are not, however, as useful for measuring dimensions of teaching such as teacher–student interaction, teacher engagement, and quality of enactment. Several studies have shown that teacher self-reports on their teaching on anonymous surveys are highly correlated with classroom observations and teacher logs and that one-time surveys that ask teachers questions about the content and strategies that they emphasize in the classroom are quite valid and reliable in measuring teachers' instruction (Mullens, 1995; Mullens & Gayler, 1999; Mullens & Kasprzyk, 1996, 1999; Schmidt et al., 1997; Shavelson, Webb, & Burstein, 1986; Smithson & Porter, 1994).

Another issue is our within-group analyses. The subgroup analysis of White, Black, low-SES, and high-SES students allows us to examine how teacher and teaching variables operate for particular groups of students.

We are careful not to compare across groups, though, which would require mean tests of betas for each set of models. Our focus is addressing the achievement gap, so within-group comparisons are useful. However, these analyses should not be used to compare findings across different racial/ethnic or income groups.

The strengths of our study include our separation of summer growth to focus on academic-year growth, which arguably has the most potential to be influenced by teacher and teaching quality variables, though lagged effects are, of course, possible. In addition, we focus on both teacher and teaching quality in a national longitudinal sample that allows measurement of growth, not just cross-sectional correlations or gains, as is more common.

RESEARCH QUESTION 1: WHAT IS THE DISTRIBUTION OF TEACHER AND TEACHING QUALITY DURING THE FIRST YEAR OF KINDERGARTEN?

We found evidence that lower achieving students are initially assigned to new teachers and to teachers who use more basic procedural approaches to instruction; in contrast, higher achieving students are initially assigned to teachers who tend to use more advanced procedural (multistep algorithms, especially advanced for kindergartners) and conceptual approaches to mathematics.

These findings occur in the context of a growth modeling analysis that found no effects on student growth of being a new teacher. The only consistent finding for type of instruction was that advanced procedural approaches were related to achievement growth. This suggests that these aspects of teacher and teaching quality may operate as sorting variables, which may explain a part of the findings of past cross-sectional and gain studies that would likely interpret correlations between teachers and teaching as part of the effect of instruction.

RESEARCH QUESTION 2: TO WHAT EXTENT DO TEACHER QUALITY, TIME SPENT ON INSTRUCTION, AND TYPE OF INSTRUCTION PREDICT GROWTH IN STUDENT ACHIEVEMENT IN KINDERGARTEN AND FIRST GRADE?

Teacher Quality. We did not find consistent or strong relationships between teacher quality and achievement growth in either kindergarten or first grade. These findings are generally consistent with Guarino et al.'s (2006) kindergarten ECLS study. Our mixed findings for certification are consistent with earlier studies (Darling-Hammond et al., 2001;

Goldhaber & Brewer, 2000; Smith et al., 2005). Alternative certification overall had a positive effect on achievement growth but a negative effect for high-SES students, and having a teacher with no certification was positively related to growth for students with low SES. Keeping post-hoc explanations to a minimum, these findings might be due to the fact that requirements for certification and paths for alternative certification vary across states and districts. For example, alternative certification in a high-need district might be a proxy for teachers with higher mathematics content knowledge than the regular workforce, whereas in a wealthy suburban district, which on average has qualified, certified mathematics teachers, alternative certification is more likely to be a proxy for having less content knowledge than the average teacher. That is, alternative certification may be bringing teachers with content-area expertise into high-need schools, which serves as an advantage for what we know is generally an underqualified population of teachers, especially in mathematics (Ingersoll, 1999). We did find that high certification, where it was significant, was always positive. High certification is measured here by the following question: "What type of teaching certification do you have: The highest certification available (permanent or long term)?" This, too, might vary across states in its meaning.

Having a teacher without a bachelor's degree in mathematics slowed growth in first grade mathematics but not in kindergarten. We expected that proxies for content knowledge, such as having a BA in mathematics, would affect achievement growth, and the results supported this hypothesis. It is unusual for teachers of the early grades to have a strong background in mathematics (NCES, 2002a). Further, teachers of the early grades sometimes argue that basic knowledge of mathematics is enough and question why deeper knowledge of mathematics is useful when teaching basic addition and subtraction (Berger, Desimone, Herman, Garet, & Margolin, 2002; Miller, Herman, Garet, Desimone, & Zhang, 2002). Previous research explains the dynamics of how a deeper understanding of elementary math and how students learn mathematics can be a powerful mechanism to enable a teacher to better diagnose and respond to student mistakes and shape activities that are productive for a range of learners (Ball, 1990; Ma, 1999). Here our findings show support for the idea that such content knowledge (at least our crude proxy) supports growth over time in first graders. Previous work in kindergarten may suggest why type of teaching would not be as influential; several studies have shown that other areas are more crucial to early cognitive development, such as parent involvement in learning (Barnett & Escobar, 1987; Booth & Dunn, 1996).

Other teacher quality measures in our analysis—professional develop-

ment in mathematics, courses taken in mathematics, and teacher's years of experience—were, for the most part, not related to achievement growth. Absent a direct measure of teachers' content and pedagogical content knowledge (Hill, Schilling, & Ball, 2004), these can all be considered proxies for teacher knowledge and skill. The extent to which any one of them is related to achievement growth might be considered an indication of how close they are to measuring the knowledge and skills related to effective teaching. We found that teacher certification and having a BA in mathematics were significant predictors, whereas the other measures were not. It should be noted that the more local measures, dependent on district context and offerings, such as experience and professional development, did not factor into growth, whereas more calibrated measures, such as certification and having a degree, did so. In addition, the two measures of teacher quality that were significant represent a combination or series of educational experiences (e.g., a BA in mathematics represents passing several courses in mathematics), whereas the other measures represent piecemeal incremental experiences (e.g., one math course, a certain number of hours in professional development). A different configuration of the incremental variables—for example, a measure of over 40 hours of professional development, or several mathematics courses—may have made a difference. Previous research suggests that there is a threshold above which teacher learning experiences must go before they are effective in changing practice (Desimone, Porter, Garet, Yoon, & Birman, 2002; Garet, Porter, Desimone, Birman, & Yoon, 2001). Further, we had no measure of the quality of the professional development that teachers experienced, which obviously limits its predictive ability.

Time spent teaching and type of instruction. There are opposing views concerning the benefits to students of teachers' relative emphasis on procedural and conceptual mathematics instruction (Loveless, 2001). Several small-scale cross-sectional and longitudinal studies have shown positive results for conceptual approaches to mathematics (e.g., Carpenter et al., 1989; Newman & Associates, 1996). Based on these findings, we expected students exposed to conceptual instruction to grow at a faster rate than students exposed predominantly to basic procedural or advanced procedural instruction. We found that of the three types of instruction, the only consistent finding across groups was that kindergarten and first grade students whose teachers emphasized advanced procedural instruction, here measured by a focus on multistep addition and subtraction, grew at a more accelerated pace than their peers. Teachers were not asked *how* they taught multistep algorithms (e.g., what cognitive demands they used in their teaching, such as memorize or estimate), so this

finding might be interpreted to mean that focusing on particular top-ics—here double-digit addition, subtraction, and division—fostered student achievement growth in mathematics. In the early grades, topic coverage is especially important because many students are dependent on only one or two teachers to give them exposure to particular content.

More powerful than distinguishing between types of teaching, we found that the number of minutes spent on mathematics instruction in first grade was associated with achievement for traditionally disadvantaged populations—Black and low-SES students. This is consistent with the idea that disadvantaged students do not have the exposure to learning opportunities at home, especially in mathematics, and so are more dependent on the school to provide these mathematics experiences.

Other studies have found stronger effects for different types of instruction (e.g., Carpenter et al., 1989; see Loveless, 2001). Usually, however, these are smaller scale longitudinal studies that use curriculum-based tests. It is not uncommon to find no effects on a standardized measure of achievement, so it is noteworthy that we did find effects on sorting and for advanced procedural and time spent on mathematics.

RESEARCH QUESTION 3: TO WHAT EXTENT DO TEACHER QUALITY, TIME SPENT ON INSTRUCTION, AND TYPE OF INSTRUCTION NARROW THE ACHIEVEMENT GAP FOR BLACK STUDENTS AND STUDENTS FROM LOW-INCOME FAMILIES?

Our calculations show that it would take 100 minutes of increased math instruction per day to close the Black–White achievement gap by 10%, and 100 minutes to narrow the SES achievement gap between the 25th and the 75th percentiles by 13%. These findings are reminiscent of the opportunity-to-learn literature of several decades ago, which emphasized the importance of increasing exposure to academic content, not necessarily details about how the content was covered. These findings are similar to Guarino et al.'s (2006) ECLS kindergarten study, which found that full-day kindergarten and minutes spent on mathematics instruction made more of a difference for student gains than did types of instruction. We did not find effects for full-day kindergarten, but this might be because the way we created our minutes spent on mathematics measure captures most of the "full day" variation.

CONCLUSION

Students spend considerably more time out of school than in school (Hofferth & Sandberg, 2001), and the achievement gaps exists before

children start school, so it is unlikely that it could be fully closed by the schools. However, school is often revered as the "great equalizer," or at least a social institution with the potential for an equalizing function, so time and effort spent exploring how schools might address inequality continues to be compelling. Here we examine particular aspects of schooling that might explain the recent finding that schools do narrow the achievement gap (Downey et al., 2004).

The findings here do not provide overwhelming support for a particular type of mathematics instruction, but rather for what has been coined "opportunity to learn"—the time spent on mathematics content. We found weak or no effects for teacher quality and type of instruction, which lends support to the hypothesis that initial allocation of students to teachers explains at least some of the correlation between teacher and teaching quality found in cross-sectional and gains studies. With a national sample and a standardized test, though, we would not expect our analyses to be especially sensitive to distinctions in types of teaching. Our most powerful finding was that for low-income students and Black students, minutes spent on instruction made a difference for achievement growth in kindergarten and first grade, though these effects are small.

Here we are encouraged that something happening in the classroom can make a difference for the achievement gap. Our intercept analysis showed that low achievers tend to get worse teachers. This is an old problem of the most disadvantaged students getting the weakest teachers. If time on instruction matters, and disadvantaged students are more likely to get the weakest teachers who spend less time on instruction, then we can identify an area where schooling may be exacerbating the achievement gap but has the potential to ameliorate it.

Notes

1. The Asian, Hispanic, Native American, and mixed race subgroups were not large enough to allow separate subgroup analysis.

2. To create the class-size dummy, we regressed dummies for class size from 15 to 35 on achievement and found that a cutoff of 27 explained the most variance. Similarly, to create the LEP dummy variable, we regressed dummy variables for each decile of LEP, such as 0%, 10%, and 20%– 100%, and found that the 10% cutoff explained the most variance.

3. In separate sets of analyses, we used multiple imputation and found that results were very close to mean imputation. For ease of interpretation, we present results for mean imputation here.

4. As explained in our Discussion section, though we use subgroup analyses, we do not compare directly across models, which would require mean tests of betas for each set of models.

References

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis, 25*, 171–191.

Allison, P. D. (2001). *Missing data.* Thousand Oaks, CA: Sage.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *Elementary School Journal, 90*, 449–466.

Ballou, D. (1996). Do public schools hire the best applicants? *Quarterly Journal of Economics, 111*, 97–133.

Barnett, W. S., & Escobar, C. M. (1987). The economics of early intervention: A review. *Review of Educational Research, 57*, 387–414.

Barr, R., Wiratchai, N., & Dreeben, R. (1983). *How schools work.* Chicago: University of Chicago Press.

Berger, A., Desimone, L., Herman, R., Garet, M., & Margolin, J. (2002). *Content of state standards and the alignment of state assessments with state standards.* Washington, DC: U.S. Department of Education.

Blomqvist, N. (1977). On the relation between change and initial value. *Journal of the American Statistical Association, 72*, 746–749.

Booth, A., & Dunn, J. F. (Eds.). (1996). *Family-school links: How do they affect educational outcomes?* Mahwah, NJ: Erlbaum.

Bourdieu, P., & Passeron, J. (1977). *Reproduction in education, society, and culture.* Beverly Hills, CA: Sage.

Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America: Educational reform and contradictions of economic life.* New York: Basic Books.

Braswell, J. S., Lutkus, A. D., Grigg, W. S., Santapau, S. L. Tay-Lim, B.S.-H., & Johnson, M. S. (2001). *The nation's report card: Mathematics 2000* (NCES 2001–517). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*, 499–531.

Carroll, J. (1963). A model of school learning. *Teachers College Record, 64*, 722–733.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Grayson, W., Trigatti, B., et al.. (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education, 22*, 3–29.

Cohen, D., & Ball, D. (1990). Policy and practice: An overview. *Educational Evaluation and Policy Analysis, 12*, 347–353.

Cohen, D., & Hill, H. (2000). *Learning policy: When state education reform works.* New Haven, CT: Yale University Press.

Cohen, D. K., & Hill, H. C. (2001). *Learning policy: When state education reform work*s. New Haven, CT: Yale University Press.

Cohen, D., McLaughlin, M., & Talbert, J. (Eds.). (1993). *Teaching for understanding: Challenges for policy and practice.* San Francisco: Jossey-Bass.

Coleman, J. S., Campbell, E. Q., Hobson, C. F., McPartland, J. M., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Office of Education, U.S. Government Printing Office.

Condron, D. J., & Roscigno, V. J. (2003). Disparities within: Unequal spending and achievement in an urban school district. *Sociology of Education, 76*, 18–36.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*, 227–268.

Cremin, L. A. (1951). *The American common school: An historical conception.* New York: Bureau of Publications, Teachers College, Columbia University.

D'Agostino, J. V. (2000). School effects on students' longitudinal reading and math achievements. *School Effectiveness and School Improvement, 11,* 197–235.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1). Retrieved from http://olam.ed.asu.edu/epaa/v8n1/

Darling-Hammond, L., Berry, B., & Thorenson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis, 23,* 57–77.

Desimone, L., Porter, A. C., Garet, M., Yoon, K., & Birman, B. (2002). Effects of professional development on teachers' instruction: Results from a three-year study. *Educational Evaluation and Policy Analysis, 24,* 81–112.

Desimone, L., Smith, T., & Frisvold, D. (2007). Is NCLB increasing teacher quality for students in poverty? In A. Gamoran (Ed.), *Will standards-based reform in education help close the poverty gap?* (pp. 89–119). Washington, DC: Brookings Institution Press.

Desimone, L. M., Smith, T. S., Hayes, S., & Frisvold, D. (2005). Beyond accountability and average math scores: Relating multiple state education policy attributes to changes in student achievement in procedural knowledge, conceptual understanding and problem solving in mathematics. *Educational Measurement: Issues and Practice, 24*(4), 5–18.

Downey, D., von Hippel, P., & Broh, B. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review, 69,* 613–635.

Ehrenberg, R. G., & Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from High School and Beyond. *Economics of Education Review, 13*(1), 1–17.

Ehrenberg, R. G., & Brewer, D. J. (1995). Did teachers' verbal ability and race matter in the 1960s? Coleman revisited. *Economics of Education Review, 14,* 1–21.

Elmore, R. F., Peterson, P. L., & McCarthy, S. J. (1996). *Restructuring in the classroom: Teaching, learning, and school organization.* San Francisco: Jossey-Bass.

Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation, 28,* 465–499.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 265–298). Washington, DC: Brookings Institution.

Fetler, M. (1999, March). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives, 7*(9). Retrieved from http://epaa.asu.edu/epaa/v7n9.html

Fryer, R. G., Jr., & Levitt, S. D. (2004). Understanding the Black–White test score gap in the first two years of school. *Review of Economics and Statistics, 86,* 447–464.

Gamoran, A. (1986). Instructional and institutional effects of ability groups. *Sociology of Education, 59,* 185–198.

Gamoran, A., & Mare, R. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement or neutrality? *American Journal of Sociology, 94,* 1146–1183.

Gamoran, A., Porter, A. C., Smithson, J., & White, P. A. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis, 19,* 325–338.

Gamoran, A., Secada, W. G., & Marrett, C. B. (2000). The organizational context of teaching and learning: Changing theoretical perspectives. In M. T. Hallinan (Ed.), *Handbook of the sociology of education* (pp. 37–63). New York: Kluwer Academic.

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? Analysis of a national sample of teachers. *American Education Research Journal, 38,* 915–945.

Geary, D. C. (2001). A Darwinian perspective on mathematics and instruction. In T. Loveless (Ed.), *The great curriculum debate: How should we teach reading and math?* (pp. 85–107). Washington, DC: Brookings Institution Press.

Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66,* 361–396.

Goldhaber, D. D., & Brewer, D. J. (1997). Evaluating the effect of teacher degree level on educational performance. In W. Fowler (Ed.), *Developments in school finance 1996* (pp. 197–210). NCES 97-535. U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

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

Guarino, C. M., Hamilton, L. S., Lockwood, J. R., Rathbun, A. H., & Hausken, E. G. (2006.) *Teacher qualification, instructional practices, and reading and mathematics gains of kindergartners* (Research and Development Report, NCES 2006-031). Washington, DC: U.S. Department of Education.

Hamilton, L. S., McCaffrey, D., Klein, S. P., Stecher, B. M., Robyn, A., & Bugliari, D. (2003). Teaching practices and student achievement: Studying classroom-based education reforms. *Educational Evaluation and Policy Analysis, 25,* 1–29.

Hart, B. H., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Paul H. Brookes.

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

Hedges, L. V., & Nowell, A. (1998). Black-White test score convergence since 1965. In C. Jencks & M. Phillips (Eds.), *The Black–White test score gap* (pp. 149–181). Washington, DC: Brookings Institution Press.

Heyns, B. (1978). *Summer learning and the effects of schooling.* New York: Academic Press.

Heyns, B. (1987). Schooling and cognitive development: Is there a season for learning? *Child Development, 58,* 1151–1160.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., et al. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher, 25*(4), 12–21.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Wearne, D., Murray, H., et al. (1997). *Making sense: Teaching and learning mathematics with understanding.* Portsmouth, NH: Heinemann.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal, 105,* 11–30.

Hofferth, S. L., & Sandberg, J. F. (2001). How American children use their time. *Journal of Marriage and Family, 62,* 295–308.

Ingersoll, R. M. (1999). *Teacher turnover, teacher shortages and the organization of schools.* Seattle: University of Washington, Center for the Study of Teaching and Policy.

Ingersoll, R. (2002, January). *Out of field teaching, educational inequality, and the organization of schools: An exploratory analysis.* Seattle: University of Washington, Center for Study of Teaching and Policy.

Knapp, M. S., & Shields, P. M. (Eds.). (1990). *Better schooling for children of poverty: Alternatives to conventional wisdom* (Vol. 2). Washington DC: U.S. Department of Education, Office of Planning, Budget, and Evaluation.

Knapp, M. S., Shields, P. M., & Turnbull, B. J. (1992). *Academic challenge for the children of poverty. Summary report.* Washington, DC: U.S. Department of Education.

Kozma, R., & Croninger, R. (1992). Technology and the fate of at-risk students. *Education and Urban Society, 24,* 440–453.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27,* 29–63.

Lee, V. E., Smith, J. B., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education, 70,* 128–150.

Levine, M. (Ed.). (1988). *Professional practice schools: Building a model* (ED 313344). Washington, DC: American Federation of Teachers.

Little, R., & Rubin, D. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley.

Loveless, T. (Ed.). (2001). *The great curriculum debate: How should we teach reading and math?* Washington, DC: Brookings Institution Press.

Lubienski, S. (2002). A closer look at Black–White mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data. *Journal of Negro Education, 71,* 269–287.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States.* Mahwah, NJ: Erlbaum.

Manswell Butty, J. (2001). Teacher instruction, student attitudes, and mathematics performance among 10th and 12th grade Black and Hispanic students. *Journal of Negro Education, 70*(1/2), 19–37.

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis, 21*(1), 29–45.

Milesi, C., & Gamoran, A. (2005, April). *Effect of class size and instruction on kindergarten achievement.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Miller, L., Herman, R., Garet, M., Desimone, L., & Zhang, Y. (2002). *State mathematics standards: Policies, instructional supports, aligned instruction, and student achievement.* Washington, DC: U.S. Department of Education.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13,* 125–145.

Monk, D. H., & King, J. R. (1994). Multi-level teacher resource effects on pupil performance in secondary mathematics and science: The role of teacher subject matter preparation. In R. Ehrenberg (Ed.), *Contemporary policy issues: Choices and consequences in education* (pp. 29–58). Ithaca, NY: ILR Press.

Mosteller, F., & Moynihan, D. P. (Eds.). (1972). *On equality of educational opportunity.* New York: Vintage Books.

Mullens, J. (1995). *Classroom instructional processes: A review of existing measurement approaches and their applicability for the teacher follow-up survey* (NCES 95-15). Washington, DC: National Center for Education Statistics.

Mullens, J. E., & Gayler, K. (1999). *Measuring classroom instructional processes: Using survey and case study field test results to improve item construction* (NCES 1999-08). Washington, DC: National Center for Education Statistics.

Mullens, J., & Kasprzyk, D. (1996). Using qualitative methods to validate quantitative survey instruments. In *1996 Proceedings of the Section on Survey Research Methods* (pp. 638–643). Alexandria, VA: American Statistical Association.

Mullens, J., & Kasprzyk, D. (1999). *Validating item responses on self-report teacher surveys.* Washington, DC: U.S. Department of Education.

Murnane, R. J., & Phillips, B. R. (1981). Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance. *Economics of Education Review, 1,* 453–465.

National Center for Education Statistics. (2000). *Early Childhood Longitudinal Study—Kindergarten base year: Data files and electronic codebook.* Washington DC: Author.

National Center for Education Statistics. (2002a). *Schools and staffing survey: 1999–2000. Overview of the data for public, private, public charter, and Bureau of Indian Affairs elementary and secondary schools* (NCES 2002-213).Washington, DC: Author.

National Center for Education Statistics. (2002b). *User's manual for the ECLS-K Longitudinal Kindergarten-First Grade Public-Use Data Files and Electronic Codebook.* Washington, DC: Author.

National Commission on Teaching and America's Future. (1996, September). *What matters most: Teaching for America's future.* New York: Author.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

Newman, F., & Associates. (1996). *Authentic achievement: Restructuring schools for intellectual quality.* San Francisco: Jossey-Bass.

Oakes, J. (1985). *Keeping track: How schools structure inequality.* New Haven, CT: Yale University Press.

Pellegrino, J. W., Baxter, G. P., & Glaser, R. (1999). Addressing the "two disciplines" problem: Linking theories of cognition and learning with assessment and instructional practice. *Review of Research in Education, 24,* 307–353.

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the Black-White test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 229–272). Washington, DC: Brookings Institution Press.

Porter, A. C. (2005). Prospects for school reform and closing the achievement gap. In C. A. Dwyer (Ed.), *Measurement and research in the accountability era* (pp. 59–95). Mahwah, NJ: Erlbaum.

Porter, A. C., Kirst, M. W., Osthoff, E. J., Smithson, J. L., & Schneider, S. A. (1993). *Reform up close: An analysis of high school mathematics and science classrooms.* Madison: University of Wisconsin–Madison.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Rock, D., & Pollack, J. (2002). *Early childhood longitudinal study—kindergarten class of 1988–99 (ECLS-K), psychometric report for kindergarten through 1st grade* (NCES 2002-005). U.S. Department of Education, NCES. Washington, DC: U.S. Government Printing Office.

Rowan, B., Correnti, R., & Miller, R. (2002). *What large-scale survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools.* Philadelphia: Consortium for Policy Research in Education.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* New York: Wiley.

Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, H. C., Wiley, D. E., Cogan, L. S., et al. (2001). *Why schools matter: A cross-national comparison of curriculum and learning.* San Francisco: Jossey-Bass.

Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1997). *A splintered vision: An investigation of U.S. science and mathematics. Executive summary.* Lansing: U.S. National Research Center for the Third International Mathematics and Science Study, Michigan State University. Retrieved from http://ustimss.msu.edu/splintrd.pdf

Shavelson, R. J., Webb, N. M., & Burstein, L. (1986). Measurement of teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 1–36). Washington, DC: American Educational Research Association.

Shouse, R. (2001). The impact of traditional and reform-style practices on student mathematics achievement. In T. Loveless (Ed.), *The great curriculum debate: How should we teach reading and math?* (pp. 108–133). Washington, DC: Brookings Institution Press.

Silver, E., & Lane, S., (1995). Can instructional reform in urban middle schools help students narrow the mathematics performance gap? Some evidence from the QUASAR project. *Research in Middle Level Education Quarterly, 18*(2), 49–70.

Singer, J. D., & Willet, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York: Oxford University Press.

Slavin, R., Madden, N., Karweit, N., Livermon, B., & Dolan, L. (1990). Success for All: First-year outcomes of a comprehensive plan for reforming urban education. *American Educational Research Journal, 27*, 255–278.

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

Smith, J., Lee, V., & Newmann, F. (2001). *Instruction and achievement in Chicago elementary schools. Improving Chicago's Schools.* Chicago: Consortium on Chicago School Research.

Smithson, J. L., & Porter, A. C. (1994). Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum—the Reform Up Close study (CPRE Research Report Series No. 31). Madison: University of Wisconsin, Consortium for Policy Research in Education.

Spillane, J., & Zeuli, J. (1999, Spring). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis, 21*(1), 1–27.

Strauss, R., & Sawyer, E. (1986). Some new evidence on teacher and student competencies. *Economics of Education Review, 5*(1), 41–48.

Tate, W. (1997). Race-ethnicity, SES, gender, and language proficiency trends in mathematics achievement: An update. *Journal for Research in Mathematics Education, 28,* 652–680.

Thomson, G. H. (1924). A formula to correct for the effect of errors of measurement on the correlation of initial values with gains. *Journal of Experimental Psychology, 7*, 321–324.

Tourangeau, K., Nord, C., Lê, T., Pollack, J. M., & Atkins-Burnett, S. (2006). *ECLS-K: Combined user's manual for the ECLS-K fifth-grade data files and electronic codebooks.* Washington, DC: U.S. Department of Education.

U.S. Department of Education. (2003). *Digest of education statistics, 2002.* Washington, DC: National Center for Education Statistics.

Wenglinsky, H. (2000). *How teaching matters: Brining the classroom back into discussions of teacher quality.* Princeton NJ: Educational Testing Service.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10*(12), 1–31. Retrieved from http://epaa.asu.edu/epaa/v10n12

Wenglinsky, H. (2004). Closing the racial achievement gap: The role of reforming instructional practices. *Education Policy Analysis Archives, 12*(64). Retrieved from http://epaa.asu.edu/epaa/v12n64/

Wiley, D., & Yoon, B. (1995). Teacher reports on opportunity to learn: Analyses of the 1993 California Learning Assessment System. *Educational Evaluation and Policy Analysis, 17*, 355–370.

Wright, S., Horn, S., & Sanders, W. (1997) Teachers and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57–67.

Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the Early Childhood Longitudinal Study—kindergarten class of 1998–99. *American Educational Research Journal, 41,* 191–229.

Appendix A. Descriptive Statistics

	Full Sample		White Sample		Black Sample		High SES		Low SES	
Level 1 Missing Covariates	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Missing SES	0.03	0.16	0.01	0.12	0.04	0.21	0	0	0	0
Missing Days Absent in Kindergarten	0.06	0.24	0.07	0.26	0.07	0.26	0.05	0.23	0.06	0.25
Missing Days Absent in First Grade	0.23	0.42	0.21	0.41	0.28	0.45	0.22	0.41	0.25	0.43
Level 2 Missing Covariates										
Kindergarten Missing Dummy Variables	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Missing Math courses	0.06	0.23	0.05	0.22	0.08	0.27	0.05	0.22	0.05	0.23
Missing Teacher Experience	0.04	0.19	0.03	0.16	0.03	0.18	0.04	0.18	0.03	0.18
Missing BA Var.	0.1	0.29	0.08	0.27	0.1	0.3	0.1	0.3	0.09	0.27
Missing Cert.Var.	0.07	0.25	0.06	0.23	0.09	0.27	0.07	0.24	0.07	0.25
Missing the Full-day variable	0.1	0.29	0.08	0.27	0.09	0.29	0.08	0.27	0.1	0.29
Missing LEP variable	0.09	0.28	0.08	0.26	0.09	0.27	0.08	0.27	0.09	0.29
Missing % minority variable	0.12	0.32	0.11	0.3	0.11	0.31	0.11	0.31	0.12	0.32
Missing Minutes of math	0.12	0.24	0.11	0.23	0.12	0.24	0.1	0.23	0.12	0.24
Missing Conceptual Variable	0.1	0.29	0.09	0.28	0.1	0.3	0.09	0.28	0.11	0.31
Missing Advanced Variable	0.16	0.36	0.14	0.34	0.17	0.37	0.14	0.35	0.16	0.37
Missing Basic Variable	0.1	0.3	0.09	0.29	0.09	0.29	0.09	0.29	0.11	0.31
Missing Professional Developemnt	0.03	0.16	0.02	0.15	0.03	0.17	0.02	0.14	0.02	0.16
First Grade Missing Dummy Variables										
Missing Math courses	0.05	0.18	0.04	0.17	0.05	0.18	0.04	0.16	0.05	0.18
Missing Teacher Experience	0.02	0.1	0.02	0.1	0.02	0.13	0.02	0.11	0.02	0.1
Missing BA Var.	0.02	0.13	0.02	0.13	0.03	0.15	0.03	0.14	0.02	0.12
Missing Cert.Var.	0.03	0.14	0.03	0.15	0.03	0.15	0.03	0.15	0.03	0.14
Missing LEP variable	0.02	0.1	0.02	0.11	0.03	0.12	0.02	0.1	0.02	0.1
Missing % minority variable	0.04	0.17	0.05	0.18	0.05	0.18	0.05	0.18	0.04	0.16
Missing Minutes of math	0.05	0.18	0.06	0.18	0.06	0.19	0.06	0.18	0.05	0.17
Missing Conceptual Variable	0	0	0	0	0	0	0	0	0	0
Missing Advanced Procedural Variable	0	0	0	0	0	0	0	0	0	0
Missing Basic Procedural Variable	0	0	0	0	0	0	0	0	0	0
% w/ Data on First Grade Teachers	0.87	0.34	0.9	0.29	0.88	0.32	0.92	0.27	0.88	0.33
Missing Professional Developemnt	0.01	0.09	0.01	0.08	0.01	0.1	0.01	0.08	0.01	0.09

Appendix B. Correlation Between Instruction Variables and Teacher Characteristics

Panel A: Kindergarten Teachers

	K Basic	K Conceptual	K Advanced	K New Teacher	K Experience	K BA	K Master's	K Certification	K No Certification	K Emergency Certification	K Alternative Certification	K Minutes Spent on mat
K Basic	1.00											
K Conceptual	0.65 0.00	1.00										
K Algorithm	0.58 0.00	0.48 0.00	1.00									
K New Teacher	0.06 0.01	0.09 0.00	0.04 0.05	1.00								
K Experience	0.06 0.00	0.02 0.26	0.02 0.47	-0.15 0.00	1.00							
K BA	0.14 0.00	0.06 0.00	0.09 0.00	-0.14 0.00	0.18 0.00	1.00						
K Master's	0.10 0.00	0.11 0.00	0.00 0.00	0.12 0.00	-0.11 0.00	-0.68 0.00	1.00					
K Certification	0.14 0.00	0.11 0.00	0.11 0.00	0.15 0.00	-0.20 0.00	-0.08 0.00	0.33 0.00	1.00				
K No Certification	0.03 0.24	-0.04 0.07	0.00 0.90	-0.12 0.00	0.13 0.00	-0.02 0.25	-0.08 0.00	-0.14 0.00	1.00			
K Emergency Certification	0.05 0.03	0.03 0.19	0.03 0.20	-0.12 0.00	0.25 0.00	0.11 0.00	-0.04 0.06	-0.31 0.00	-0.03 0.12	1.00		
K Certification Alternative	0.04 0.04	0.03 0.14	0.05 0.02	-0.02 0.43	0.01 0.55	0.02 0.38	-0.01 0.70	-0.13 0.00	-0.01 0.51	-0.03 0.14	1.00	
K Minutes Spent on Math Instruction	0.45 0.00	0.32 0.00	0.32 0.00	0.06 0.00	0.01 0.79	0.12 0.00	0.12 0.00	0.15 0.00	-0.04 0.06	0.00 0.83	0.04 0.10	1.00
	KtcertHI	KtcertNO	KtcertEM	KtcertAL	Kminute							

KtcertHI

KtcertNO

KtcertEM

KtcertAL

Kminute

Appendix B. Correlation Between Instruction Variables and Teacher Characteristics (continued)

Panel B: First Grade Teachers

	F Basic	F Conceptual	F Advanced	F New Teacher	F Experience	F BA	F Master's	F Certification	F No Certification	F Emergency Certification	F Alternative Certification	F Minutes Spent on mat
F Basic	1.00											
F Conceptual	0.74 / 0.00	1.00										
F Advanced	0.87 / 0.00	0.72 / 0.00	1.00									
F New Teacher	0.03 / 0.14	0.09 / 0.00	0.03 / 0.17	1.00								
F Experience	0.14 / 0.00	0.12 / 0.00	0.12 / 0.00	-0.14 / 0.00	1.00							
F BA	0.19 / 0.00	0.11 / 0.00	0.20 / 0.00	-0.13 / 0.00	0.22 / 0.00	1.00						
F Master's	0.35 / 0.00	0.38 / 0.00	0.32 / 0.00	0.12 / 0.00	-0.07 / 0.00	-0.64 / 0.00	1.00					
F Certification	0.07 / 0.00	0.12 / 0.00	0.05 / 0.01	0.07 / 0.00	-0.05 / 0.02	-0.13 / 0.00	0.21 / 0.00	1.00				
F No Certification	0.05 / 0.04	0.02 / 0.32	0.04 / 0.04	-0.08 / 0.00	0.02 / 0.00	0.00 / 0.04	0.00 / 0.11	-0.05 / 0.03	1.00			
F Emergency Certification	0.07 / 0.00	0.04 / 0.10	0.06 / 0.01	-0.10 / 0.00	0.20 / 0.26	0.04 / 0.06	-0.03 / -0.02	-0.08 / 0.00	-0.04 / 0.06	1.00		
F Certification Alternative	0.08 / 0.00	0.03 / 0.12	0.06 / 0.01	-0.09 / 0.00	0.06 / 0.00	0.11 / 0.00	0.41 / -0.06	0.00 / -0.04	0.01 / 0.61	-0.03 / 0.23	1.00	
F Minutes Spent on Math Instruction	0.62 / 0.00	0.58 / 0.00	0.63 / 0.00	0.06 / 0.00	0.08 / 0.00	0.18 / 0.00	0.41 / 0.00	0.09 / 0.00	0.01 / 0.71	0.08 / 0.00	0.04 / 0.08	1.00
FtcertHI												
FtcertNO												
FtcertEM												
FtcertAL												
Fminute												

Appendix C. Predicted Growth Curves for Different Racial Groups, SES Percentiles, Minutes of Math Instruction, and Amounts of Advanced Math Instruction

Panel A: Math Achievement by Race

Days since K. Began	Race					
	All	White	Black	Hispanic	Asian	American Indian
1	16.88	17.38	15.96	15.52	18.04	14.91
275	28.95	29.7	26.8	27.29	30.55	27.71
365	31.05	31.72	29.01	29.62	33.33	29.05
639	44.4	45.27	41.63	43.36	45.25	42

Panel B: Math Achievement by Socio-Economic Status

Days since K. Began	Socio-Economic Status				
	bottom 1%	25th percentile	50th percentile	75th percentile	top 1%
1	12.87	15.54	16.67	18.15	22.09
275	24.22	27.36	28.7	30.44	35.08
365	26.03	29.37	30.79	32.64	37.56
639	40.04	42.94	44.17	45.77	50.04

Panel C: Math Achievement by Minutes of Math Instruction

Days since K. Began	Minutes of Math Instruction per Day				
	Average Minutes	Zero Minutes	Avg. Minus 100 Minutes	Avg. Min. + 100 Minutes	Avg. Min. + Three Hours
1	16.88	16.92	16.9	16.86	16.85
275	28.95	28.73	28.83	29.07	29.17
365	31.05	30.83	30.93	31.17	31.27
639	44.4	43.6	44.03	44.76	45.06

Panel D: Math Achievement by Days of Advanced Procedural Instruction

Days since K. Began	Days of Advanced Procedural Instruction per Month				
	Avg. Days	Zero Days	Avg. Minus Four Days	Avg. Plus Four Days	Avg. Plus Ten Days
1	16.88	16.52	16.56	17.21	17.69
275	28.95	28.19	28.28	29.62	30.62
365	31.05	30.29	30.38	31.72	32.72
639	44.4	43.09	43.45	45.35	46.77

Appendix D. Predicted Growth Curves for Differential Increases in Minutes of Math Instruction and Amount of Advanced Procedural Instruction for the Different Subsamples

Panel A: Math Achievement for White and Black Samples

Race

Days since K. Began	Black	White	Black + 10 Days More of Advanced Instruction Per Month	Black + 3 Hours More Minuites of Instruction Per Day	White + 10 Days More of Advanced Instruction Per Month	White + 3 Hours More Minuites of Instruction Per Day
1	14.77	17.88	15.47	15.13	18.98	17.91
275	24.58	29.88	26.38	24.94	32.08	29.91
365	26.38	31.71	28.18	26.74	33.91	31.74
639	39.41	43.93	40.65	40.75	46.67	43.96

Panel B: Math Achievement for the Different SES Samples

Socio-Economic Status

Days since K. Began	Highest SES Quartile	Lowest SES Quartile	Highest SES Quartile + 10 Days More of Advanced Instruction Per Month	Highest SES Quartile + 3 Hours More Minuites of Instruction Per Day	Lowest SES Quartile + 10 Days More of Advanced Instruction Per Month	Lowest SES Quartile + 3 Hours More Minuites of Instruction Per Day
1	20.92	13.09	22.42	20.56	13.49	13.27
275	32.58	24.21	34.35	32.71	25.71	24.89
365	34.49	25.85	36.26	34.62	27.35	26.53
v639	47.73	38.55	49.51	47.87	40.59	40.21

LAURA M. DESIMONE is associate professor, Graduate School of Education, University of Pennsylvania. Her research focuses on the effects of policies on teaching and learning, policy implementation, and improving surveys for use in policy research. She studies teacher's professional development, induction and mentoring, accountability, and teacher quality issues. Recent publications include: Desimone, L., Smith, T., & Frisvold, D. (in press). How similar are student and teacher reports of classroom instruction? *Educational Policy*; and Desimone, L. (in press). How can we best measure teacher's professional development and its effects on teachers and students? *Educational Researcher*.

DANIEL LONG is assistant professor of sociology at Wesleyan University. His research focuses on sociology of education and social stratification. He has researched trends in educational attainment among different racial ethnic/groups in the United States; researched how school and teacher characteristics influence the growth of student achievement in elementary school; and examined international achievement tests to estimate the effects of private schools and school resources. His recent publications include "Equality of Educational Opportunity: A 40-Year Retrospective" with Adam Gamoran; and "Whither the Virtuous Cycle: Past and Future Trends in Black–White Inequality in Educational Attainment" with Adam Gamoran and Sean Kelly.