Education Policy Analysis Archives

Volume 8 Number 1 **January 1, 2000** **ISSN 1068-2341**

A peer-reviewed scholarly electronic journal
Editor: Gene V Glass, College of Education
Arizona State University

Articles appearing in **EPAA** are abstracted in the *Current Index to Journals in Education* by the ERIC Clearinghouse on Assessment and Evaluation and are permanently archived in *Resources in Education*.

Teacher Quality and Student Achievement: A Review of State Policy Evidence

Linda Darling-Hammond
Stanford University

Abstract
Using data from a 50-state survey of policies, state case study analyses, the 1993-94 Schools and Staffing Surveys (SASS), and the National Assessment of Educational Progress (NAEP), this study examines the ways in which teacher qualifications and other school inputs are related to student achievement across states. The findings of both the qualitative and quantitative analyses suggest that policy investments in the quality of teachers may be related to improvements in student performance. Quantitative analyses indicate that measures of teacher preparation and certification are by far the strongest correlates of student achievement in reading and mathematics, both before and after controlling for student poverty and language status. State policy surveys and case study data are used to evaluate policies that influence the overall level of teacher qualifications within and across states. This analysis suggests that policies adopted by states regarding teacher education, licensing, hiring, and professional development may make an important difference in the qualifications and capacities that teachers bring to their work. The

implications for state efforts to enhance quality and equity in public education are discussed. (Note 1)

Introduction

For many years, educators and researchers have debated which school variables influence student achievement. As policymakers become more involved in school reform, this question takes on new importance since their many initiatives rely on presumed relationships between various education-related factors and learning outcomes. Some research has suggested that "schools bring little influence to bear upon a child's achievement that is independent of his background and general social context" (Coleman et al., 1966, p. 325; see also Jencks et al., 1972). Other evidence suggests that factors like class size (Glass et al., 1982; Mosteller, 1995), teacher qualifications (Ferguson, 1991), school size (Haller, 1993), and other school variables may play an important role in what students learn.

As new standards for student learning have been introduced across the states, greater attention has been given to the role that teacher quality plays in student achievement (National Commission on Teaching and America's Future, 1996; National Education Goals Panel, 1998). In the last few years, more than 25 states have enacted legislation to improve teacher recruitment, education, certification, or professional development (Darling-Hammond, 1997a). While some evidence suggests that better qualified teachers may make a difference for student learning at the classroom, school, and district levels, there has been little inquiry into the effects on achievement that may be associated with large-scale policies and institutional practices that affect the overall level of teachers' knowledge and skills in a state or region. This paper reports on one such study, which combines state level case studies and quantitative analyses of state-level achievement data to examine whether and how state policies may influence teachers' capabilities and student learning.

Using data from a 50-state policy survey conducted by the National Commission on Teaching and America's Future, case studies of selected states conducted under the auspices of the Center for the Study of Teaching and Policy, the 1993-94 Schools and Staffing Surveys (SASS), and the National Assessment of Educational Progress (NAEP) sponsored by the National Center for Education Statistics, the study examines the ways in which teacher qualifications and other school inputs, such as class size, are related to student achievement across states, taking student characteristics into account. In addition, these data and state case study data are used to evaluate policies that influence the overall level of teacher qualifications within and across states.

Previous Research

Despite conventional wisdom that school inputs make little difference in student learning, a growing body of research suggests that schools can make a difference, and a substantial portion of that difference is attributable to teachers. Recent studies of teacher effects at the classroom level using the Tennessee Value-Added Assessment System and a similar data base in Dallas, Texas, have found that differential teacher effectiveness is a strong determinant of differences in student learning, far outweighing the effects of differences in class size and heterogenity (Sanders & Rivers, 1996; Wright, Horn, & Sanders, 1997; Jordan, Mendro, & Weerasinghe, 1997). Students who are assigned to several ineffective teachers in a row have significantly lower achievement and gains in achievement than those who are assigned to several highly effective teachers in sequence (Sanders & Rivers, 1996). Teacher effects appear to be additive and cumulative, and

generally not compensatory. These studies also find troubling indicators for educational equity, noting evidence of strong bias in assignment of students to teachers of different effectiveness levels (Jordan, Mendro, & Weerasinghe, 1997), including indications that African American students are nearly twice as likely to be assigned to the most ineffective teachers and half as likely to be assigned to the most effective teachers (Sanders & Rivers, 1996). These studies did not, however, examine the characteristics or practices of more and less effective teachers.

These issues have been the topic of much other research over the last 50 years. Variables presumed to be indicative of teachers' competence which have been examined for their relationship to student learning include measures of academic ability, years of education, years of teaching experience, measures of subject matter and teaching knowledge, certification status, and teaching behaviors in the classroom. The results of these studies have been mixed; however, some trends have emerged in recent years.

General Academic Ability and Intelligence While studies as long ago as the 1940s have found positive correlations between teaching performance and measures of teachers' intelligence (usually measured by IQ) or general academic ability (Hellfritsch, 1945; LaDuke, 1945; Rostker, 1945; Skinner, 1947), most relationships are small and statistically insignificant. Two reviews of such studies concluded that there is little or no relationship between teachers' measured intelligence and their students' achievement (Schalock, 1979; Soar, Medley, & Coker, 1983). Explanations for the lack of strong relationship between measures of IQ and teacher effectiveness have included the lack of variability among teachers in this measure and its tenuous relationship to actual performance (Vernon, 1965; Murnane, 1985). However, other studies have suggested that teachers' verbal ability is related to student achievement (e.g., Bowles & Levin, 1968; Coleman et al., 1966; Hanushek, 1971), and that this relationship may be differentially strong for teachers of different types of students (Summers & Wolfe, 1975). Verbal ability, it is hypothesized, may be a more sensitive measure of teachers' abilities to convey ideas in clear and convincing ways (Murnane, 1985).

Subject Matter Knowledge Subject matter knowledge is another variable that one might think could be related to teacher effectiveness. While there is some support for this assumption, the findings are not as strong and consistent as one might suppose. Studies of teachers' scores on the subject matter tests of the National Teacher Examinations (NTE) have found no consistent relationship between this measure of subject matter knowledge and teacher performance as measured by student outcomes or supervisory ratings. Most studies show small, statistically insignificant relationships, both positive and negative (Andrews, Blackmon & Mackey, 1980; Ayers & Qualls, 1979; Haney, Madaus, & Kreitzer, 1986; Quirk, Witten, & Weinberg, 1973; Summers & Wolfe, 1975).

Byrne (1983) summarized the results of thirty studies relating teachers' subject matter knowledge to student achievement. The teacher knowledge measures were either a subject knowledge test (standardized or researcher-constructed) or number of college courses taken within the subject area. The results of these studies were mixed, with 17 showing a positive relationship and 14 showing no relationship. However, many of the "no relationship" studies, Byrne noted, had so little variability in the teacher knowledge measure that insignificant findings were almost inevitable. Ashton and Crocker (1987) found only 5 of 14 studies they reviewed exhibited a positive relationship between measures of subject matter knowledge and teacher performance.

It may be that these results are mixed because subject matter knowledge is a positive influence up to some level of basic competence in the subject but is less important thereafter. For example, a controlled study of middle school mathematics

teachers, matched by years of experience and school setting, found that students of fully certified mathematics teachers experienced significantly larger gains in achievement than those taught by teachers not certified in mathematics. The differences in student gains were greater for algebra classes than general mathematics (Hawk, Coble, & Swanson, 1985). However, Begle and Geeslin (1972) found in a review of mathematics teaching that the absolute number of course credits in mathematics was not linearly related to teacher performance.

It makes sense that knowledge of the material to be taught is essential to good teaching, but also that returns to subject matter expertise would grow smaller beyond some minimal essential level which exceeds the demands of the curriculum being taught. This interpretation is supported by Monk's (1994) more recent study of mathematics and science achievement. Using data on 2,829 students from the Longitudinal Study of American Youth, Monk (1994) found that teachers' content preparation, as measured by coursework in the subject field, is positively related to student achievement in mathematics and science but that the relationship is curvilinear, with diminishing returns to student achievement of teachers' subject matter courses above a threshold level (e.g., five courses in mathematics).

In a multilevel analysis of the same data set, Monk and King (1994) found both positive and negative, generally insignificant effects of teachers' subject matter preparation on student achievement. They did find some evidence of cumulative effects of prior as well as proximate teachers' subject matter preparation on student performance in mathematics. Effects differed for high- and low-achieving students and for different grade levels. In a review of 65 studies of science teachers' characteristics and behaviors, Druva and Anderson (1983) found students' science achievement was positively related to the teachers' course taking background in both education and in science. The relationship between teachers' training in science and student achievement was greater in higher level science courses, a result similar to that found by Hawk, Coble, and Swanson (1985) in mathematics.

It may also be that the measure of subject matter knowledge makes a difference in the findings. Measures of course-taking in a subject area have more frequently been found to be related to teacher performance than have scores on tests of subject matter knowledge. This might be because tests necessarily capture a narrower slice of any domain. Furthermore, in the United States, most teacher tests have used multiple-choice measures that are not very useful for assessing teachers' ability to analyze and apply knowledge. More authentic measures may capture more of the influence of subject matter knowledge on student learning. For example, a test of French language teachers' speaking skill was found to have significant correlation to students' achievement in speaking and listening (Carroll, 1975).

Despite concerns that education majors may be less well prepared in their subject areas than are academic majors (Galambos, 1985), comparisons of teachers with degrees in education vs. those with degrees in disciplinary fields have found no relationship between degree type and teacher performance (Murnane, 1985). This may be because certification requirements reduce the variability in course backgrounds found for teachers with different degree types. For example, many states require the equivalent of an academic major or minor in the field to be taught as part of the education degree for high school teachers, regardless of the department granting the degree (NASDTEC, 1997). Given the standardizing influences of licensing requirements within states but substantial differences in licensing requirements across states, within-state studies are likely to find less variation in teachers' education backgrounds than might be found in cross-state studies.

Knowledge of Teaching and Learning Studies have found a somewhat stronger and more consistently positive influence of education coursework on teachers' effectiveness. Ashton and Crocker (1987) found significant positive relationships between education coursework and teacher performance in four of seven studies they reviewed—a larger share than those showing subject matter relationships. Evertson, Hawley, and Zlotnik (1985) reported a consistent positive effect of teachers' formal education training on supervisory ratings and student learning, with 11 of 13 studies showing greater effectiveness for fully prepared and certified vs. uncertified or provisionally certified teachers. With respect to subject matter coursework, 5 of 8 studies they reviewed found no relationship and the other three found small associations.

Reviewing findings of the National Longitudinal Study of Mathematical Abilities, Begle (1979) found that the number of credits a teacher had in mathematics methods courses was a stronger correlate of student performance than was the number of credits in mathematics courses or other indicators of preparation. Similarly, Monk's (1994) study of student's mathematics and science achievement found that teacher education coursework had a positive effect on student learning and was sometimes more influential than additional subject matter preparation. In an analysis of science teaching, Perkes (1967- 68) found that teachers' coursework credits in science were not significantly related to student learning, but coursework in science education was significantly related to students' achievement on tasks requiring problem solving and applications of science knowledge. Teachers with greater training in science teaching were more likely to use laboratory techniques and discussions and to emphasize conceptual applications of ideas, while those with less education training placed more emphasis on memorization.

In a study of more than 200 graduates of a single teacher education program, Ferguson and Womack (1993) examined the influences on 13 dimensions of teaching performance of education and subject matter coursework, NTE subject matter test scores, and GPA in the student's major. They found that the amount of education coursework completed by teachers explained more than four times the variance in teacher performance (16.5 percent) than did measures of content knowledge (NTE scores and GPA in the major), which explained less than 4 percent. In a similar study which compared relative influences of different kinds of knowledge on 12 dimensions of teacher performance for more than 270 teachers, Guyton and Farokhi (1987) found consistent strong, positive relationships between teacher education coursework performance and teacher performance in the classroom as measured through a standardized observation instrument, while relationships between classroom performance and subject matter test scores were positive but insignificant and relationships between classroom performance and basic skill scores were almost nonexistent. Another program-based study by Denton and Lacina (1984) found positive relationships between the extent of teachers' professional education coursework and their teaching performance, including their students' achievement.

It may be that the positive effects of subject matter knowledge are augmented or offset by knowledge of how to teach the subject to various kinds of students. That is, the degree of pedagogical skill may interact with subject matter knowledge to bolster or reduce teacher performance. As Byrne (1983) suggested:

> It is surely plausible to suggest that insofar as a teacher's knowledge provides the basis for his or her effectiveness, the most relevant knowledge will be that which concerns the particular topic being taught and the relevant pedagogical strategies for teaching it to the particular types of pupils to whom it will be taught. If the teacher is to teach fractions, then it is

> knowledge of fractions and perhaps of closely associated topics which is of major importance.... Similarly, knowledge of teaching strategies relevant to teaching fractions will be important. (p. 14)

The kind and quality of in-service professional development as well as pre-service education may make a difference in developing this knowledge. Several recent studies have found that higher levels of student achievement are associated with mathematics teachers' opportunities to participate in sustained professional development grounded in content-specific pedagogy linked to the new curriculum they are learning to teach (Cohen & Hill, 1997; Wiley & Yoon, 1995; Brown, Smith, & Stein, 1995). In these studies, both the kind and extent of professional development mattered for teaching practice and for student achievement.

The National Assessment of Educational Progress has also documented how specific kinds of teacher learning opportunities correlate with their students' reading achievement. On average, in the 1992 and 1994 assessments, 4th grade students of teachers who were fully certified, who had master's degrees, and who had had professional coursework in literature-based instruction did better than other students on reading assessments (NCES, 1994; NCES, n.d.). While these relationships were modest, the relationships between specific teaching practices and student achievement were often quite pronounced, and these practices were in turn related to teacher learning opportunities. NAEP analyses found that teachers who had had more professional training were more likely to use teaching practices that are associated with higher reading achievement on the NAEP tests--use of trade books and literature, integration of reading and writing, and frequent visits to the library--and were less likely to engage in extensive of use of reading kits, basal readers, workbooks, and multiple choice tests for assessing reading, practices that the NAEP analyses found to be associated with lower levels of student achievement. Interestingly, students of teachers who had had more training in phonics instruction did noticeably less well than other students in both years. Often, this kind of training, narrowly cast, is focused heavily on the use of basal readers and workbooks rather than an integrated approach that teaches decoding skills in the context of other important reading skills and language development strategies.

Other studies have found that students achieve at higher levels and are less likely to drop out when they are taught by teachers with certification in their teaching field, by those with master's degrees, and by those enrolled in graduate studies (Council for School Performance, 1997; Knoblock, 1986; Sanders, Skonie-Hardin, & Phelps, 1994). However, like the NAEP analyses described above, these are simple correlational analyses that do not take into account other school resources or student characteristics like poverty or language background that may affect student performance.

Continuity of teachers' learning may also matter. In earlier work, Hanushek (1971) demonstrated that the *recency* of voluntary educational experience was related to teacher performance. Penick and Yager (1983) found that teachers in exemplary science programs had higher levels of education and more recent educational experiences than others, even though they were older than the average science teacher. As Murnane (1985) suggests, these findings may indicate that it is not only the knowledge acquired with ongoing professional development (which may represent more recent advances in the knowledge base) but also the teacher's enthusiasm for learning that relates to increased student achievement.

Teaching Experience Other studies of the effects of teacher experience on student learning have found a relationship between teachers' effectiveness and their years

of experience (Murnane & Phillips, 1981; Klitgaard & Hall, 1974), but not always a significant one or an entirely linear one. While many studies have established that inexperienced teachers (those with less than three years of experience) are typically less effective than more senior teachers, the benefits of experience appear to level off after about five years, especially in non-collegial work settings (Rosenholtz, 1986). A possible cause of this curvilinear trend in experience effects is that older teachers do not always continue to grow and learn and may grow tired in their jobs. Furthermore, the benefits of experience may interact with educational opportunities. Veteran teachers in settings that emphasize continual learning and collaboration continue to improve their performance (Rosenholtz, 1984). Similarly, very well-prepared beginning teachers can be highly effective. For example, some recent studies of 5-year teacher education programs--programs that include a bachelor's degree in the discipline and master's in education as well as a year-long student teaching placement--have found graduates to be more confident than graduates of 4-year programs and as effective as more senior teachers (Andrew & Schwab, 1995; Denton & Peters, 1988).

It is also possible that uneven effects of experience in cross-sectional studies can be the result of cohort effects (for example, cohorts of teachers hired in times of shortage may be less well-qualified than those hired when schools can be more selective) or of attrition effects (for example, disproportionate early attrition of more able teachers may leave a less capable senior force on average) (Murnane & Phillips, 198; Vance & Schlechty, 1982). Presumably, the direction of this effect would change if retention policies kept the most able beginning teachers in the profession. Since experience is also correlated with teacher education and certification status, these variables may be confounded in some analyses.

Certification Status Certification or licensing status is a measure of teacher qualifications that combines aspects of knowledge about subject matter and about teaching and learning. Its meaning varies across the states because of differences in licensing requirements, but a standard certificate generally means that a teacher has been prepared in a state-approved teacher education program at the undergraduate or graduate level and has completed either a major or a minor in the field(s) to be taught plus anywhere from 18 to 40 education credits, depending on the state and the certificate area, including between 8 and 18 weeks of student teaching. (The norm is about 30 education credits and about 12 to 15 weeks of student teaching.) There are only a few states that have requirements outside these parameters; however, individual teacher education programs often require more preparation than the state demands in education, in clinical practice, and in the content area(s) to be taught. Most states now also require one or more tests of basic skills, subject matter knowledge, and/or teaching knowledge or skills as the basis for the initial or continuing license or for admission to teacher education.

While most states have been increasing their standards since the 1980s, more than 30 states still allow the hiring of teachers who have not met their licensing standards, a practice that has been on the increase in some states as demand has grown in recent years. Some allow the hiring of teachers with no license. Others issue emergency, temporary, or provisional licenses to candidates who, depending on the state, may or may not have met varying requirements (e.g., a bachelors degree, a certificate in another teaching field, a basic skills test). More than 40 states have also initiated alternate route provisions for candidates who enter through postbaccalaureate programs. Most of these are master's degree programs which offer an education degree that meets all of the normal state requirements but does so in a fashion tailored to individuals, like mid-career entrants, who already have a bachelor's degree. Some states allow candidates

to complete a short summer course of study and assume full teaching responsibilities, with or without completing additional coursework.

In times of relatively low demand, like most of the 1980s, virtually all teachers were certified and there was too little variability to find effects of this variable in large-scale studies. Most studies of the influence of training and certification on teacher performance are from the high-demand era of the 1960s and 1970s and from the 1990s when demand increased again. Studies in different subject matter fields that compare teachers with and without preparation have typically found higher ratings and greater student learning gains for teachers who have more formal preparation for teaching. In addition to the studies of science and mathematics teachers cited earlier, these include reading and elementary education (Hice, 1970; LuPone, 1961; McNeil, 1974), early childhood education (Roupp et al., 1979), gifted education (Hansen, 1988), and vocational education (Erekson and Barr, 1985). In a review of research, Evertson, Hawley, and Zlotnik (1985) concluded:

> (T)he available research suggests that among students who become teachers, those enrolled in formal preservice preparation programs are more likely to be effective than those who do not have such training. Moreover, almost all well planned and executed efforts within teacher preparation programs to teach students specific knowledge or skills seem to succeed, at least in the short run (p.8).

Other studies point out the differences in the perceptions and practices of teachers with differing amounts and kinds of preparation. A number of studies suggest that the typical problems of beginning teachers are lessened for those who have had adequate preparation prior to entry (Adams, Hutchinson, & Martray, 1980; Glassberg, 1980; Taylor & Dale, 1971). Studies of teachers admitted with less than full preparation--with no teacher preparation or through very short alternate routes--have found that such recruits tend to be less satisfied with their training (Darling-Hammond, Hudson, & Kirby, 1987; Jelmberg, 1995), and they tend to have greater difficulties planning curriculum, teaching, managing the classroom, and diagnosing students' learning needs (Bents & Bents, 1990; Darling-Hammond, 1992; Lenk, 1989; Feiman-Nemser & Parker, 1990; Gomez & Grobe, 1990; Grady, Collins, & Grady, 1991; Grossman, 1989; Mitchell, 1987; National Center for Research on Teacher Learning, 1992; Rottenberg & Berliner, 1990). Principals, supervisors, and colleagues tend to rate them less highly on their instructional skills (Bents & Bents, 1990; Jelmberg, 1995; Lenk, 1989; Feiman-Nemser & Parker, 1990; Gomez & Grobe, 1990; Mitchell, 1987; Texas Education Agency, 1993), and they tend to leave teaching at higher-than- average rates (Darling-Hammond, 1992; Lutz & Hutton, 1989; Stoddart, 1992).

These findings are reflected in Gomez and Grobe's (1990) study of the performance of alternate route candidates in Dallas, who receive a few weeks of summer training from the district before they assume full teaching responsibilities. Although these candidates were rated near the average on some aspects of teaching, they were rated lower on such factors as their knowledge of instructional techniques and instructional models. The performance of alternate route candidates was also much more uneven than that of trained teachers, with a much greater proportion of them--from 2 to 16 times as many--rated "poor" on each of the teaching factors evaluated. The strongest effects of this unevenness were seen in students' achievement in language arts, where the achievement gains of students of alternate route teachers, adjusted for initial student scores, were significantly lower than those of students of traditionally trained teachers.

Two studies of alternative certification in Texas have reportedly failed to find such gaps in the performance of students of alternative and traditionally licensed teachers (cited in Goldhaber & Brewer, 1999). A study of Houston's alternative certification program by Goebel, Romacher, and Sanchez (1989) reported no evidence of differential student outcomes and little evidence of teacher effects. However, this study did not control for students' initial test scores and did not match comparison teachers by years of experience. First year traditionally trained teachers were compared to two groups of alternative certification recruits, one with 1-4 years of experience and the other with 5-7 years of experience. Thus, this study did not include adequate controls to allow measurement of effects. Another study by Barnes, Salmon, and Wale (1989) reported second-hand that two districts reported equivalent outcomes for alternative and traditional program teachers but did not present any empirical data or discussion of methodology. The study's table listing program types evaluated included 1 to 2-year university-based master's programs (which are called "alternative" in Texas because they are not undergraduate models) as well as district alternative programs that generally offer only a few weeks of summer training. In this case, the "alternative" group included programs providing extensive graduate level training along with those with very little preparation, thus preventing assessment of the effects of preparation on teacher effectiveness. With non-comparable groups and no controls, it is impossible to draw inferences from either of these studies.

Some recent multivariate studies of student achievement at the school and district level have found a substantial influence of teachers' qualifications on what students learn, especially when scores on licensing examinations are included. In an analysis of nearly 900 Texas school districts that evaluated the effects of many school input variables and controlled for student background and district characteristics, Ronald Ferguson (1991) found that combined measures of teachers' expertise--scores on a licensing examination, master's degrees, and experience--accounted for more of the inter-district variation in students' reading and mathematics achievement (and achievement gains) in grades 1 through 11 than student socioeconomic status. An additional, smaller contribution to student achievement was made by lower pupil-teacher ratios and smaller schools in the elementary grades.

Of the teacher qualifications variables, the strongest relationship was found for scores on the state licensing examination, a test that measures both basic skills and teaching knowledge. The effects were so strong, and the variations in teacher expertise so great, that after controlling for socioeconomic status, the large disparities in achievement between black and white students were almost entirely accounted for by differences in the qualifications of their teachers. Ferguson also found that every additional dollar spent on more highly qualified teachers netted greater increases in student achievement than did less instructionally focused uses of school resources.

Another study (Strauss & Sawyer, 1986) found that North Carolina's teachers' average scores on the National Teacher Examinations (a licensing test which measures subject matter and teaching knowledge) had a strong influence on average school district test performance. Taking into account per-capita income, student race, district capital assets, student plans to attend college, and pupil/teacher ratios, teachers' test scores had a strikingly large effect on students' failure rates on the state competency examinations: a 1% increase in teacher quality (as measured by NTE scores) was associated with a 3 to 5% decline in the percentage of students failing the exam. The authors' conclusion is similar to Ferguson's:

Differences in State Policies Regarding Teaching

Despite logical presumptions and research evidence that student learning depends substantially on what teachers know and can do, states differ greatly in the extent to which they invest in teachers' learning as a key policy lever. At the front end of the career, there is wide variation in the standards to which entering teachers and teacher education institutions are held. Licensing standards are noticeably different from state to state, as are state commitments to enforcing these standards. Later access to professional development is also widely disparate.

In high-standards states like Wisconsin or Minnesota, for example, a prospective high school teacher must complete a bachelor's degree that includes a full major in the subject area to be taught plus coursework covering learning theory, child and adolescent development, subject matter teaching methods, curriculum, effective teaching strategies, uses of technology, classroom management, behavior and motivation, human relations, and the education of students with special needs. In the course of this work, the teacher must complete at least 18 weeks of student teaching in Wisconsin (at least a college semester in Minnesota) under the supervision of a cooperating teacher who meets minimum standards. In Minnesota, this experience must include work in a multicultural setting and with special needs students. If teachers are asked to teach outside the field of their major for part of the day, they must already be licensed with at least a minor in that field, and can receive a temporary license in the new field only briefly while completing a major. By contrast, in Louisiana, prospective high school teachers can be licensed without even a minor in the field they will be teaching. The state does not require them to have studied curriculum, teaching strategies, classroom management, uses of technology, or the needs of special education students, and they can receive a license with only six weeks of student teaching (NASDTEC, 1997; Darling-Hammond, 1997a).

In addition to differences in the standards themselves, there are great differences in the extent to which they are enforced. Whereas some states do not allow districts to hire unqualified teachers, others routinely allow the hiring of candidates who have not met their standards, even when qualified teachers are available. In Wisconsin and eleven other states, for example, no new elementary or secondary teachers were hired without a license in their field in 1994. By contrast, in Louisiana, 31% of new entrants were unlicensed and another 15% were hired on substandard licenses. At least six other states allowed 20% or more of new public school teachers to be hired without a license in their field (Darling-Hammond, 1997a, Appendix A). Studies of teacher hiring show that even when there are an adequate number of qualified teachers in the labor market--which was the case nationally and in most states from the early 1980s through the mid-1990s--some districts hire unlicensed teachers because of cumbersome and poorly managed hiring procedures that discourage qualified entrants, perennially late hiring (e.g. waiting until late August or September to hire), patronage hiring, preferences for hiring lower salaried staff, and inequalities in salary schedules caused by state funding formulas and by local decisions to use budgets for purposes other than teacher salaries (see e.g. Haberman, 1995; Johanson and Gips, 1992; Pflaum and Abramson, 1990; National Commission on Teaching and America's Future; Wise, Darling-Hammond, and Berry, 1987).

More than 30 states allow teachers to be hired on temporary or emergency licenses without having completed preparation or having met other licensing requirements. During the late 1980s and early 1990s, at least 50,000 emergency or substandard licenses were issued annually by states (NCTAF, 1996). Nationally, in 1994, 27% of those who were new entrants into public school teaching held no license or a substandard license in their main teaching field (Darling-Hammond, 1997a). Even the rigor of these restricted

licenses varies. States such as Minnesota will issue a restricted license only to a teacher who has already been fully prepared in a teaching field but who needs to complete additional coursework in order to enter from out-of-state or switch to a new field or teaching level. Such a license is only good for one year, while the necessary coursework is completed. Others, including Louisiana, will issue an emergency license to a person who does not even hold a bachelor's degree and will allow it to be renewed for several years while the candidate makes little progress toward becoming licensed.

It is certainly true that differences in student enrollment growth, coupled with teacher production rates and attrition, construct different levels of teacher demand that can affect the ease or difficulty of hiring within states. While incentives to enter and stay in teaching are affected by policies governing salaries, working conditions, and teacher education funding, student enrollments are less amenable to policy control. It is reasonable to ask whether these differences in operational teaching standards are mostly a function of demographic trends beyond the control of state policymakers. In examining state variations in hiring practices, however, it is clear that a number of high-growth states have enacted and maintained high standards for entry to teaching while many low-growth states have not. Policies appear to be at least as important as demographics in determining the qualifications of teachers hired and retained.

Because of these differences in licensing standards and enforcement, in 1994, more than 80% of high school teachers of academic courses in Wisconsin and Minnesota had fully met state certification requirements and had at least a college major in the field they teach. Four other states--Connecticut, Iowa, Montana, and North Dakota--reported similarly well-qualified teaching forces in that year. The comparable proportion of teachers with full state certification and a major in their field in Louisiana was only 64%. (An additional six states had fewer than two-thirds of their teachers similarly prepared.)

Interestingly, students in Minnesota and Wisconsin have typically scored at the top of the distribution on national assessments of reading and mathematics, along with the four other states who share similarly well-qualified teachers. Together these states held six of the top ten spots in the national rankings in reading and mathematics in 1994 and 1996. Students in Louisiana have typically scored near the bottom of the NAEP distributions--no higher than 47th of 51 states in any of the assessments reported by 1996. The other six states with similar proportions of teachers holding a license and a major in their field all fall in the bottom quartile of states in the national rankings of average student achievement scores (Campbell et al., 1996; Darling-Hammond, 1997a, pp. 13, 26; Reese et al., 1997). Some have quipped that state-level student achievement in the U.S. can be best predicted by proximity to Canada--which in turn may be a proxy for variations among states in factors ranging from demographics (e.g., student poverty, parent education, and race) to political culture and spending on education. The distributions of scores described above could indeed partly support the "Canada hypothesis," which I test below.

States also differ greatly in the levels of funding they allocate to preservice and inservice teacher education, in the standards they apply to teacher education institutions and to schools, in the types and extent of professional learning opportunities and the incentives for professional study they make available to educators, and the extent to which they require or fund induction supports for beginning teachers. To illustrate these differences, in 1997 only three states required professional accreditation for schools of education and only nine funded induction programs that provided a structured program of mentoring for beginning teachers, including trained, state-funded mentors. Student teaching requirements ranged from 5 weeks in Massachusetts to 18 weeks in Wisconsin. As of 1994, the proportions of academic high school teachers teaching with both a

license and a major in their field ranged from a low of 52% to a high of 85% across states. The proportions of mathematics teachers teaching with less than a minor in the field ranged from a low of 9% to a high of 56% (Darling-Hammond, 1997a, Appendices A and B). This means that a student in one state might have only one chance in ten of being taught by an out-of-field teacher, while a student in another state might have more than a 50% chance of being taught a subject by a teacher who is not adequately prepared in that subject.

In every category of possible investment in teachers' knowledge and in every area in which standards for teaching are set (e.g., licensing, accreditation, advanced certification, on- the-job evaluation), there are substantial differences in the policies and practices employed by states. States with some of the highest, most consistently enforced standards for teachers have tended to cluster in the upper Midwest (Minnesota, Wisconsin, Iowa, Nebraska, North Dakota, Missouri, Montana, Kansas). States with the lowest and least well-enforced standards have tended to include many in the southeast (Louisiana, Mississippi, Georgia, South Carolina) and in remote locations (Alaska, Hawaii). Some states have developed relatively ambitious standards for teaching but do not enforce them for large numbers of candidates (California, New York). Others have made major investments in preservice and inservice teacher development in recent years that have affected a substantial share of the teaching force (e.g., Connecticut, Kentucky, North Carolina, West Virginia). The possible outcomes of these cross-state differences are discussed below.

Trends in Student Achievement: Policy Hypotheses

In their book, *The Manufactured Crisis*, Berliner and Biddle (1995) noted that while U.S. secondary school students tend to score below the median in international assessments of mathematics and science, students in some states score as high as those in the top-ranked countries in the world while students in others score among the bottom-ranked. U.S. students also perform relatively better in some fields than others. For example, U.S. students have compared favorably with students in other countries in reading and at about the median in general science. However, in mathematics and physical science, U.S. students do much more poorly: In the most recent international assessments, 8th graders ranked 18th out of 25 countries that met the TIMSS guidelines in mathematics and 17th out of 25 countries in physics. Twelfth graders did even more poorly (Darling-Hammond, 1997a, pp. 28-29).

Although it may be purely coincidental, these differences in rankings are similar to the differences in teacher qualifications across these fields. Since the early 1980s, the U.S. has made major investments in teacher preparation in the area of reading. Not only are almost all elementary school teachers fully certified (more than 95%), an increasing number have been prepared in programs that have a strong emphasis on training to teach reading; there has also been a large increase in the number of reading specialists throughout the 1980s. In general science and biology, where U.S. middle and high school students scored at about the median on the most recent international assessments, there are relatively few uncertified or out-of-field secondary teachers (about 18% of the total). By contrast, in mathematics and physical science, where U.S. students fall well below the international norms, teacher qualifications are much weaker. In addition to the fact that most U.S. elementary teachers have had little background in mathematics, about 30% of U.S. mathematics teachers and 50% of physical science teachers at the high school level have been teaching with less than a minor in the field, many of them uncertified (Darling-Hammond, 1997, p.28 and Appendix Table 3). While these are only

casual observations, other evidence point in similar directions.

Long-term Achievement Trends by State

Not only do U.S. students appear to perform least well in the fields in which U.S. teachers are least well prepared, the states that repeatedly lead the nation in student achievement in mathematics and reading have among the most highly qualified teachers in the country and have made longstanding investments in the quality of teaching (see Figures 1-3). The three long-time leaders--Minnesota, North Dakota, and Iowa--have all had a long history of professional teacher policy and are among the 12 states that have state professional standards boards which have enacted high standards for persons entering the teaching profession. They are recently joined at the top of the achievement distribution by Wisconsin, Maine, and Montana, states that have also enacted rigorous standards for teaching and that are among the few which rarely hire unqualified teachers on substandard licenses. Iowa, Minnesota, Montana, North Dakota, and Wisconsin have among the lowest rates of out-of-field teaching in the country and among the highest proportions of teachers holding both certification and a major in the field they teach. (Note 2) Maine joined these states in requiring certification plus a disciplinary major when it revised its licensing standards in 1988.

These states have also been leaders in redefining teacher education and licensing. Minnesota was the first state to develop performance-based standards for licensing teachers and approving schools of education during the mid-1980s and has developed a beginning teacher mentoring program in the years since (for details, see Darling-Hammond, Wise, & Klein, 1995). Wisconsin was one of the first states to require high school teachers to earn a major in their subject area in addition to completing extensive coursework in a teacher preparation program. Thus, teacher education in Wisconsin is typically a four-and-a- half to five year process. Maine, Wisconsin, Iowa, and Minnesota have all incorporated the rigorous new standards developed by the Interstate New Teacher Assessment and Support Consortium (INTASC) (Note 3) into their licensing standards and have encouraged universities to pilot performance-based assessments of teaching using these standards.

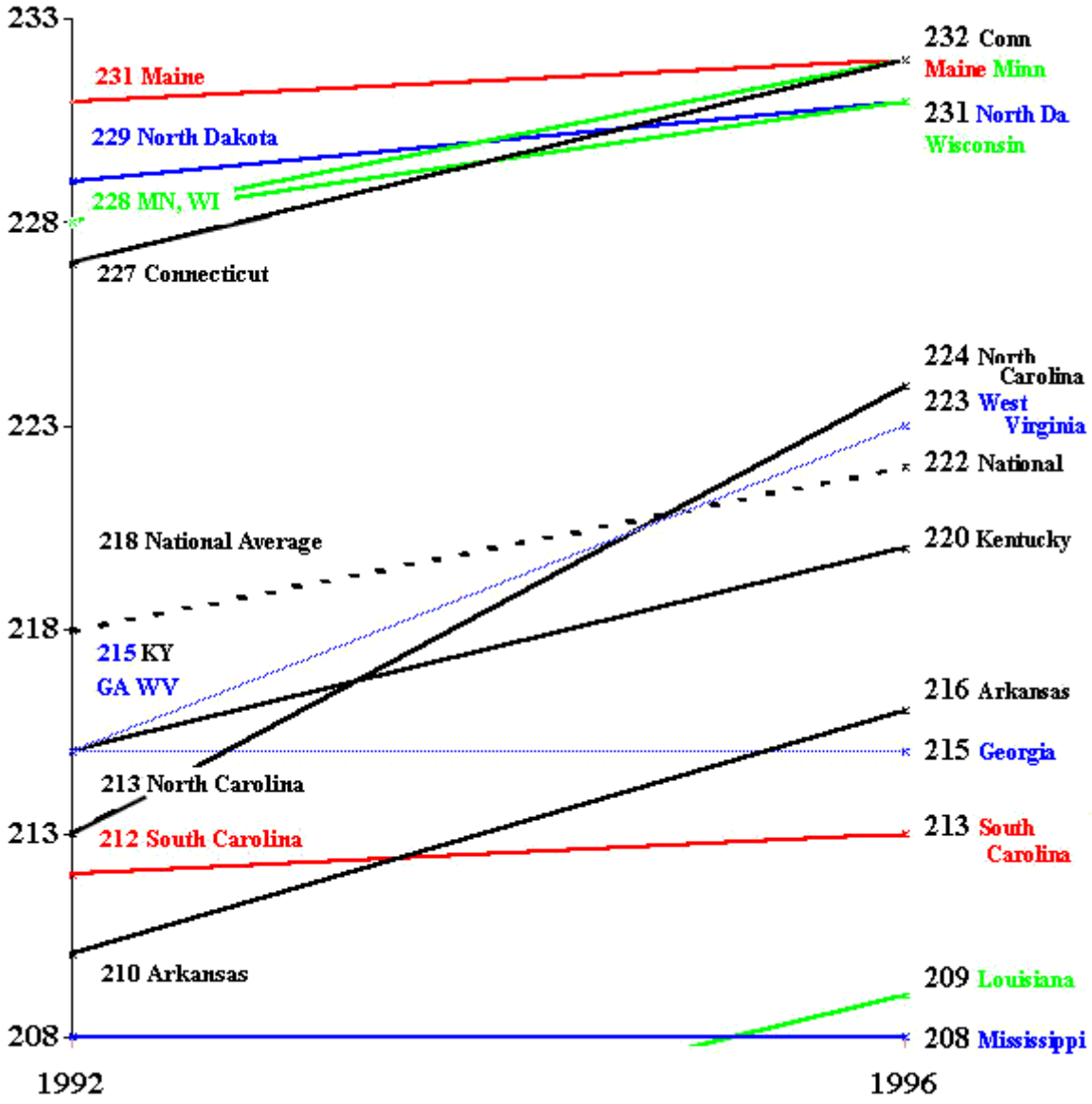

Figure 1. State Trends in Mathematics Achievement, Grade 4 (NAEP scores, 1992-1996)

Source: National Center for Education Statistics, *NAEP 1996 Mathematics Report Card for the Nation and the States,* Table 2.2, p. 28.

One can still wonder whether policies are the source of these states' strong student outcomes or whether the "Canada effect" (general education spending combined with low rates of student poverty) is responsible. Among these six states, four spent below the per pupil national average in current expenditures in 1995, and the other two spent just above the average. All, however, spent a larger percentage of their expenditures on instruction than the national average. While these states did have a lower proportion of low-income students than the national average, none fell near the tail of the distribution. There were at least twelve states with lower proportions of low-income students who scored less well on the NAEP than any of these states. However, the relative contribution of student population characteristics and school inputs is an important one to pursue further. That question is raised again below.

State Achievement Gains

Another important question is whether investments in teaching could raise achievement in states that do not have a long history of this sort. Over the last decade of reform, a few states undertook major initiatives aimed at improving the quality of teaching. From a survey of state policies, we identified five states that enacted unusually comprehensive reforms of teaching during the late 1980s and 1990s: Connecticut and North Carolina enacted the most ambitious teacher legislation of any states nationally, followed by Arkansas, Kentucky, and West Virginia, which also initiated multi-faceted reforms of teacher preparation, licensing, professional development, and compensation, accompanied by substantial investments in teacher learning.

Of the 50 states, North Carolina and Connecticut undertook the most substantial and systemic investments in teaching during the mid-1980s. Both of these states, which share relatively large high-poverty student populations, coupled major statewide increases in teacher salaries and improvements in teacher salary equity with intensive recruitment efforts and initiatives to improve preservice teacher education, licensing, beginning teacher mentoring, and ongoing professional development. Since then, North Carolina has posted the largest student achievement gains in mathematics and reading of any state in the nation, now scoring well above the national average in 4th grade reading and mathematics, although it entered the 1990s near the bottom of the state rankings. Connecticut has also posted significant gains, becoming one of the top scoring states in the nation in mathematics and reading (ranked first at the 4th grade level in mathematics and reading and in the top five at the 8th grade level), despite an increase in the proportion of low-income and limited English proficient students during that time.

North Carolina's reforms, launched with omnibus legislation in 1983, did many things simultaneously: (a) boosted salaries in the mid-1980s and again in the 1990s, (b) created a career development program that rewarded teachers for greater education and for achieving National Board Certification, (c) launched an aggressive fellowship program to recruit hundreds of able high school students into teacher preparation each year by entirely subsidizing their college education, (d) required schools of education to become professionally accredited by the National Council for the Accreditation of Teacher Education (NCATE), (e) increased licensing requirements for teachers and principals, (f) invested in improvements in teacher education curriculum, (g) created professional development academies and a North Carolina Center for the Advancement of Teaching, (h) developed teacher development networks like the National Writing Project and an analogous set of professional development initiatives in mathematics, (i) launched a beginning teacher mentoring program, and (j) introduced the most

wide-ranging set of incentives in the nation for teachers to pursue National Board certification. North Carolina now boasts more Board-certified teachers than any other state. The state was recognized in the recent National Education Goals Panel report (NEGP, 1998) for having made among the greatest gains in teacher mentoring of beginning teachers as well as the greatest achievement gains for students.

These extensive investments in teaching occurred alongside sizable investments in early childhood education and general K-12 spending increases which lowered pupil/teacher ratios slightly. In the early 1990s, new curriculum standards were introduced and accompanied by an extensive program of professional development for teachers statewide. In 1993, the state enacted an assessment system linked to the curriculum standards and substantially aligned to the NAEP tests. This assessment program, which was implemented in 1994-95, occurred too late to account for most of the gains in achievement. Its effects would require several years to appear, but it may have had some modest influence on the gains after 1994.

A recent analysis of student achievement gains on the National Assessment of Educational Progress (Grissmer & Flanagan, 1998) attributed much of the NAEP score increase in North Carolina between 1990 and 1996 to the test-based accountability system. However, the new standards and assessments were not on-line until 1995, and the rewards and sanctions component of the accountability system was not enacted until 1997, so it was clearly not a factor in these trends. Grissmer and Flanagan also note the state's large-scale investments during the 1980s in early childhood education, reduced class sizes, teacher salary increases, teacher education upgrades, and extensive professional development. All of these factors could have influenced the achievement gains observed during this time period.

North Carolina's 1997 Educational Excellence Act furthered efforts to upgrade the quality of teacher preparation and teaching quality, pouring hundreds of millions of dollars into a new set of reforms. The Act created a professional standards board for teaching and required that all colleges of education create professional development school partnerships to provide the sites for year-long student teaching practicums. It also funded a more intensive beginning teacher mentoring program, further upgraded licensing standards, created pay incentives for teachers who pursue master's degrees and National Board certification, and authorized funds to raise teacher salaries to the national average. It will be useful to watch future trends in the state.

Connecticut's strategies were similar. The state's 1986 Educational Enhancement Act spent over $300 million to boost minimum beginning teacher salaries in an equalizing fashion that made it possible for low-wealth districts to compete in the market for qualified teachers. At the same time, the state raised licensing standards by requiring a major in the discipline to be taught plus extensive knowledge of teaching and learning as part of preparation; instituted performance-based examinations in subject matter and knowledge of teaching as a basis for receiving a license; created a state-funded mentoring program which supported trained mentors for beginning teachers in their first year on the job; and created a sophisticated assessment program using state-trained assessors to determine which first-year teachers could continue in teaching. An analysis of the outcomes of this initiative found that it eliminated teacher shortages and emergency hiring, even in the cities, and created surpluses of teachers within three years of its passage (Connecticut State Department of Education, 1991).

Connecticut also required teachers to earn a master's degree in education for a continuing license and supported new, content- based professional development strategies in universities and school districts. In a National Education Goals Panel (1998) report highlighting Connecticut's strong performance and large gains in mathematics,

state officials pointed to the salary increases and teacher education investments as central to their progress. These investments include an intensive professional development program in mathematics, science, and technology which, since 1983, has offered 4-week institutes with follow-up support to elementary, middle, and high school teachers.

The state has more recently invested in new curriculum frameworks and a statewide assessment system for students using extended performance tasks and constructed response items intended to measure higher order thinking and performance skills. Launched in 1995, this system, which is tied to statewide reporting of scores and substantial new professional development, may support future gains in student achievement. In addition, the state has further extended its performance-based teacher licensing system to incorporate the new INTASC standards and to develop portfolio assessments modeled on those of the National Board for Professional Teaching Standards (NBPTS). The new teacher assessments, which are tightly linked to the student standards, require beginning teachers to demonstrate that they can implement content-based teaching standards within their subject matter field and can analyze student work and learning. Finally, as part of ongoing teacher education reforms, the state agency is supporting the creation of professional development schools linked to local universities as sites for clinical training of entering teachers.

The Connecticut and North Carolina reforms both featured substantial investments in pre-service and in-service education for teachers linked to standards that incorporate much of the current knowledge base about teaching and learning (those of NBPTS, INTASC, and/or NCATE). While the reforms also included salary increases, the dollars were linked to improved quality via heightened licensing standards. Both states sought to increase not only the quality of preparation for teachers, but also the consistency with which they enforced their standards, sharply reducing the hiring of unlicensed and under-prepared staff.

Kentucky also realized substantial achievement gains during the 1990s, after undertaking perhaps the most extensive systemic education reforms of any state in the 1980s. These included major equalization of school funding along with large increases in teacher salaries and overall spending; changes in school organization, including multi-age primary grade classrooms; investments in early childhood education; the introduction of standards and curriculum frameworks, along with portfolios and performance assessments. Changes in teacher education and licensing accompanied these reforms, including the adoption of the INTASC licensing standards developed by a consortium of more than 30 states, the introduction of new licensing tests and teacher education requirements, incentives for colleges of education to meet national professional accreditation standards; and massive investments in professional development.

All of these efforts undoubtedly combined to produce the steep gains in achievement experienced in Kentucky. By 1994, data from the Schools and Staffing Surveys showed that Kentucky teachers were much better prepared in terms of their content and teaching coursework background than in 1988 and had experienced more extensive professional development than teachers in any other state (Darling-Hammond, 1997a). A recent survey of Kentucky teachers also found that more than 80% of beginners who graduated from Kentucky colleges of education felt well-prepared for virtually all aspects of their jobs (Kentucky Institute for Educational Research, 1997), in contrast to reports about teacher education from previous studies elsewhere. Although somewhat less ambitious in their reforms, Arkansas and West Virginia also raised teacher salaries and licensing requirements and required national accreditation of

education schools during the late 1980s or early 1990s, while investing in more professional development for in-service teachers. These states also realized steeper gains in student achievement than the national average.

In a recent report, Grissmer and Flanagan (1998) focused on Texas and North Carolina for their large gains in average student achievement. They attributed Texas' gains primarily to the state's accountability system, although they also mention its shifts of resources to more disadvantaged students through school finance equalization, class size reductions, and the creation of full day kindergarten. The school funding investments that occurred in the 1980s and were continued into the following decade may indeed have made some difference in Texas students' achievement in the 1990s. However, the state's new assessment and accountability system was not initiated until 1994 and not fully implemented until 1995-96, so it could not have accounted for gains between 1990 and 1996.

Texas was not included in the above analysis of state test score gains because it was not one of the states that made large comprehensive investments in teaching during the 1980s. (Texas did make some noteworthy investments in teacher salaries and professional development in the 1990s.) In addition, however, there are questions about the stability of scores in Texas and the extent to which the posted gains are real. First, Texas included fewer than 45% of its students with disabilities in the testing pool, a much smaller share than most states (NCES, 1997, Table D3). Excessive exclusions of low-scoring students from the testing pool can cause gain scores to appear much larger than they would otherwise be. In addition, recent studies in Texas have raised concerns that much of the ostensible gain registered by African American and Latino students has been a function of grade retentions and dropouts or pushouts, which have increased substantially in recent years. These practices also make average test scores look higher by eliminating lower scoring students from the testing pool (Haney, 1999; Kurtz, 1999; Mexican American Legal Defense and Education Fund, 1999). Assuming that some of the gains in Texas are not spurious, however, it is worth noting that, in addition to the equalization of funding and investments in kindergarten and reduced class sizes, Texas was among the few states recognized by the National Education Goals Panel (1998) for large gains since the early 1990s in the proportion of beginning teachers receiving mentoring from expert veterans. Texas has also had a growing number of 5-year teacher education programs in response to an earlier reform eliminating teacher education majors at the undergraduate level.

State reform strategies during the 1980s that did not include substantial efforts to improve the nature and quality of classroom work have shown little success in raising student achievement, especially if the reforms relied primarily on student testing rather than investments in teaching. For example, the first two states to organize their reforms around new student testing systems were Georgia, with its Quality Basic Education Act (QBE) of 1985, and South Carolina, with its Education Improvement Act of 1984. These states developed extensive testing systems coupled with rewards and sanctions for students, teachers, and schools. Although both states also mandated tests for teachers, they did not link these assessments to emerging knowledge about teaching or to new learning standards, nor did they invest in improving schools of education or ongoing professional development. Few districts in either state require teachers to hold a degree in the field to be taught and full state certification as a condition of hiring. As Figures 1-3 show, student achievement in mathematics has been flat in these states while achievement in reading has declined. Since 1996, Georgia has launched an ambitious series of reforms through its P-16 Council to upgrade the quality of teacher preparation and professional development and to raise licensing standards, as well as to recruit high

ability students to teaching. Future analyses might examine whether these moves have made a difference.

There are competing hypotheses that could explain these across-state differences in achievement trajectories. One could speculate that student testing and curriculum changes are not in themselves powerful enough reforms to overcome the depressing effects on teaching quality of low standards for teacher education, licensing, and hiring, and the resulting large numbers of under-prepared teachers. On the other hand, one can argue that variables like student poverty and language background, rather than conditions that might influence the quality of teaching, are the determining factors in student achievement and that the critical differences between high- and low-achieving states are differences in their student populations.

It is interesting to compare the student achievement levels and trajectories for some of these states in comparison to geographically proximate states with similar student populations that have taken very different approaches to teaching policy. While the comparisons in Table 1 are only suggestive, they demonstrate that student achievement cannot be assumed to be only or primarily a function of demographics. Although the states that have aggressively pursued investments in teacher knowledge and skills have equal or higher levels of student poverty than nearby states that pursued other, distinctively different reform strategies, their students now achieve at higher levels. Even though all of these states increased teacher salaries during the 1990s, those that insisted on higher standards for teacher education and licensing realized gains that were not realized by states that maintained or lowered their standards for entering teaching.

Table 1
State teacher salaries, student poverty, and student achievement
NAEP 4th grade mathematics scores, 1996

	NAEP Score, 1996	Gain from 1992	% of students in poverty	Teacher Salaries Minimum	Maximum
Connecticut	232	+ 5	18.6	$28,195	$56,189
New Jersey	227	+ 0	14.6	$28,424	$58,208
North Carolina	224	+11	18.4	$20,077	$38,733
Georgia	215	+ 0	18.5	$20,065	$42,134
West Virginia	223	+ 8	22.0	$21,466	$36,378
Virginia	223	+ 2	12.6	$23,098	$38,328

Data on student achievement and poverty status from NAEP 1996 Mathematics Report Card for the States, Washington, D.C.: U.S. Department of Education, 1997, pp. 28, 139. Data on teachers' salaries from NCES, America's Teachers: Profile of a Profession, 1993-94, Washington, D.C.: U.S. Department of Education, 1996, Table A6.2.

For example, with their industrialized urban areas and affluent suburbs, Connecticut and New Jersey are demographically and economically similar states, although Connecticut has noticeably higher rates of student poverty. Despite a more affluent student population, New Jersey's students did less well than those in Connecticut on the NAEP 4th grade mathematics assessments in 1996, and, in contrast to Connecticut's students, they have not improved in recent years. Whereas Connecticut raised teachers' salaries and equalized districts' abilities to pay for qualified teachers, New Jersey decreased its requirements for teacher preparation and licensing at the end of the 1980s, reducing the amount of education coursework for entry into teaching to a maximum of 18 undergraduate credit hours and encouraging the more extensive hiring of alternative certification candidates prepared in a short summer program. These less-prepared teachers are primarily hired in low-wealth city school districts that have had radically lower revenues and salary schedules than other parts of the state.

While New Jersey's average teachers' salaries are the highest in the country, even higher than Connecticut's, New Jersey's salary increases were not tied to improvements in the qualifications of teachers or to equalization in districts' ability to pay for qualified teachers. New Jersey also lacks the rigorous licensing examinations, requirements for a major in the field and a masters in education, and state-funded mentoring for beginning teachers that Connecticut enacted in 1986. Compared to Connecticut, New Jersey has much lower rates of beginning teachers receiving mentoring and induction, much lower proportions of districts insisting on rigorous hiring standards, much lower proportions of teachers receiving professional development, much lower rates of teachers holding full certification plus a major in the field, and much higher rates of out-of-field teaching in every subject matter field (Appendix B, Tables 1-5, Darling-Hammond, 1997a).

In the same fashion, North Carolina's students now perform substantially better on the NAEP assessments than those in demographically similar Georgia, which North Carolina lagged behind in 1990. Although the states raised salaries during the 1980s and early '90s to comparable levels, Georgia did not raise standards for teacher preparation and licensing or invest heavily in teacher development at the same time. While North Carolina increased both the education and subject matter requirements for teacher preparation, introduced rigorous teacher examinations for licensing, and required national accreditation for all of its education schools during the 1980s, Georgia did little to increase expectations for either preservice or inservice preparation during those years. In addition to having had more extensive training to meet certification standards, North Carolina teachers are much more likely than their peers in Georgia to have had mentoring as beginning teachers and professional development opportunities as veterans.

And very poor West Virginia now ranks as well in elementary mathematics as its neighbor Virginia, whose students are much more affluent. Virginia, with its higher cost of living, pays its teachers more. However, West Virginia's efforts to raise salaries were accompanied by efforts to improve teacher education and licensing standards. All of West Virginia's teacher education programs must now meet national accreditation standards--a much higher set of requirements than those in Virginia, which lowered standards for education programs and licensing during the 1980s to among the lowest in the country. Like New Jersey, Virginia reduced the requirements for coursework on teaching and learning in undergraduate programs, while West Virginia raised its standards. West Virginia introduced an ambitious program of professional development even before it launched its new curriculum frameworks in the mid-1990s, and enacted a mentoring program for beginning teachers. Despite its relative wealth, Virginia hires many more unlicensed new teachers than West Virginia and its districts are less likely to

insist on rigorous hiring standards.

These kinds of contrasts can be seen in many comparisons of geographically proximate, demographically similar states that have taken different approaches to the issue of teacher investments over the last decade. Policies that jointly raise salaries and standards may offer particularly high leverage on teaching quality. It is interesting to note that, like states that introduced testing without making investments in teaching, those that have raised salaries alone, without raising standards for preparation and licensing or investing in professional development, seem not to have realized the benefits of improved student outcomes. While interesting, these observations of individual state cases could be idiosyncratic. An important question is whether similar patterns exist when viewed from a national perspective.

A National View of Teacher Qualifications and Student Achievement

To examine further the relative contributions of teaching policies and student characteristics to student achievement, this analysis uses data on public school teacher qualifications and other school inputs available from the 1993-94 Schools and Staffing Surveys (SASS) and data on student achievement and student characteristics from the 1990, 1992, 1994, and 1996 assessments in reading and mathematics administered by the National Assessment of Educational Progress. These data are the basis for regression analyses of school resource variables on student achievement scores to examine whether teacher quality indicators, as well as other school inputs, are related to student achievement at the state level, after controlling for such student characteristics as poverty and language background.

The Database The 1993-94 SASS database includes linked surveys of 65,000 teachers (52,000 public and 13,000 private); 13,000 school principals (9,500 public and 3,500 private); and 5,600 school districts. SASS is designed to provide reliable estimates of the characteristics of schools and educators at the national and state levels. It also includes information from individual teachers, school principals, and districts about salaries and compensation policies, induction policies, school climate and context variables (e.g., time to work with other teachers, teacher involvement in decision-making), professional development support, teachers' views of teaching, and their plans to remain in the profession. These analyses use the following data derived from the public school teachers' questionnaire: data on teachers' qualifications (teachers' degrees, majors, certification status), teaching assignments, and average class size. Also included in the analysis are data from the public school district questionnaire on district hiring policies (whether districts require, as a condition of hiring, full certification, graduation from an approved teacher education program, or a college major or minor in the field to be taught) and salary schedules (minimum and maximum salaries) as reported by district officials. Salary schedule data are more appropriate for gauging attractions to teaching than average salary data, which do not control for differential levels of experience and education across states. All of the SASS data were aggregated to the state level.

Teacher quality variables constructed from the SASS data include the proportion of "well-qualified teachers," defined as the proportion holding state certification and the equivalent of a major (either an undergraduate major or masters degree) in the field taught. For elementary teachers, the equivalent of a major is an elementary education degree for generalists who teach multiple subjects to the same group of students or a degree in the field taught for specialists (e.g. reading, mathematics or mathematics education, special education). The proportion of teachers who are "fully certified"

includes teachers with standard or regular certification and new teachers on probationary certificates who have completed all requirements for a license except for the completion of the probationary period (usually 2 or 3 years of beginning teaching). The proportion of teachers who are "less than fully certified" includes teachers with no certificate and those with provisional, temporary, or emergency certification.

Additional data on each state, including policies regarding teacher education and licensing (number of weeks of student teaching required, presence of a professional standards board, percentage of teacher education institutions that are NCATE accredited), were collected directly from states and professional associations (see Darling-Hammond, 1997a, Appendix A). State school spending data (current per pupil expenditures) are from the Common Core of Data (NCES, 1995).

Data from the National Assessment of Educational Progress (NAEP) include state average achievement scores for students in mathematics at the 4th grade level in 1990 and 1996 and at the 8th grade level in 1992 and 1996, as well as data on state average achievement scores for students in reading at the 4th grade level in 1992 and 1994 (Campbell, Donahue, Reese, & Phillips, 1996) and student poverty rates (Reese, Miller, Mazzeo, & Dossey, 1997).

Limitations There are a number of limitations that pertain to the data set and the analyses. First, the NAEP data derive from tests that do not measure all of the valued outcomes of schooling held by parents, teachers, and schools. They cannot represent everything that schools do or should do. In addition, state scores and changes in average scores on these measures are sensitive to differences in the population of students taking the tests, including decisions about which students will be excluded from testing and differences across states in the extent to which populations are represented in school (as a function of school- age population characteristics, dropout rates and patterns, and other variables).

Finally, the level of aggregation necessarily influences the interpretations of results. Aggregating data to the state level produces different results than one would find if one looked at similar kinds of data at the individual student, teacher, school, or district level. The direction of the differences cannot be predicted with certainty (Ferguson and Ladd, 1996). While, on one hand, the possibility of greater variability or noise exists in disaggregated analyses, it is possible that omitted variables may bias the coefficients of school input variables upward when the data are aggregated to the district or state level (Hanushek, Rivkin, and Taylor, 1995). Although the results of more and less aggregated specifications can be consistent (for example, Ferguson and Ladd's (1996) Alabama analysis found comparable influences of teacher quality and class sizes on student achievement when measured at the student and the district levels), this may not always occur. In particular, the size of relationships found between variables measured at the state level cannot be assumed to represent the effect sizes one would find in a classroom level analysis. For the purposes of assessing broad policy influences at the state level, it is nonetheless reasonable to examine state-level data as a gauge of major trends when other confirming and disconfirming evidence is available to supplement the analysis.

The Findings All analyses include public schools and teachers only. Although the sample includes all states participating in state NAEP and thus is not a representative sample from which one would draw population inferences, I report p-values as an aid to readers who wish to use them to interpret the relative sizes of relationships and the probabilities of a Type I error. Before constructing the multivariate analyses, initial bivariate correlations of school resource variables and student demographic variables with state average student test scores were conducted to examine the relationships

among variables and to select variables for inclusion in the multivariate equations. These analyses confirmed several findings reported elsewhere:

- *Student characteristics such as poverty, non-English language status, and minority status are negatively correlated with student outcomes, and usually significantly so*. These student characteristics are also significantly and negatively correlated with the qualifications of teachers; that is the less socially advantaged the students, the less likely teachers are to hold full certification and a degree in their field and the more likely they are to have entered teaching without certification.
- *Student characteristics are generally not significantly correlated with state per-pupil spending or with teachers' salary schedules*, with the exception that salary schedules are higher in states with larger percentages of minority and LEP (limited English proficient) students. Salary levels show an insignificant, negative relationship with levels of student poverty.
- *Teacher quality characteristics such as certification status and degree in the field to be taught are very significantly and positively correlated with student outcomes*. Characteristics such as education level (percentage of teachers with master's degrees) show positive but less strong relationships with education outcomes.
- *Per pupil spending (measured as current expenditures) shows a significant positive relationship with student outcomes in 4th grade reading in both years, but no relationship with student outcomes in mathematics*. This may be because the spending measure incorporates resources spent not only on teacher salaries and professional development but also on class sizes and other resources that may especially support students in the early grades as they are learning to read. Although salaries and spending are strongly related to one another ($p < .01$), teacher salary levels, unadjusted for cost of living differences, are not correlated with student outcomes when aggregated to the state level.
- *Other school resources, such as pupil-teacher ratios, class sizes, and the proportion of all school staff who are teachers, show very weak and rarely significant relationships to student achievement when they are aggregated to the state level*.

Partial correlations confirm a strong, significant relationship of teacher quality variables to student achievement even after controlling for student poverty and for student language background (LEP status) in (see Table 2 and Figure 4). The most consistent highly significant predictor of student achievement in reading and mathematics in each year tested is the proportion of well-qualified teachers in a state: those with full certification and a major in the field they teach (r between .61 and .80, $p<.001$). The strongest, consistently negative predictors of student achievement, also significant in almost all cases, are the proportions of new teachers who are uncertified (r between -.40 and -.63, $p<.05$) and the proportions of teachers who hold less than a minor in the field they teach (r between -.33 and -.56, $p<.05$). General spending and salary variables, along with class sizes, are not significantly related to achievement once student characteristics are taken into account. It should be noted, however, that this analysis did not take into account cost-of-living differentials that may affect both salaries and spending levels; controlling for such differentials could produce a different set of results with respect to these variables.

Table 2
Partial Correlations (controlling for student poverty) between Selected Teacher Quality Variables and Student Achievement on the National Assessment of Educational Progress

	Grade 4 Math, 1992	Grade 4 Math, 1996	Grade 8 Math, 1990	Grade 8 Math, 1996	Grade 4 Reading, 1992	Grade 4 Reading, 1994
% of teachers well-qualified (with full certification and a major in their field)	.71***	.61***	.75***	.67***	.80***	.75***
% of teachers out of field (with less than a minor in the field they teach)	-.48**	-.44**	-.32	-.42**	-.56**	-.33*
% of all teachers fully certified	.36*	.20	.38*	.28	.57***	.41*
% of all teachers less than fully certified	-.36*	-.23	-.33*	-.28	-.55***	-.50*
% of new entrants to teaching who are uncertified (excluding transfers)	-.51**	-.39*	.43**	-.38*	-.44**	-.47**
% of all newly hired teachers uncertified	-.40**	-.41**	-.53***	-.49**	-.59***	-.63***
Per pupil spending	.32	.28	.19	.29	.24	.27
Pupil: teacher ratio	.03	.22	.09	.12	.08	.08
Class size	-.03	.21	-.04	-.00	.08	.13

*p<.10 **p<.05 ***p<.01

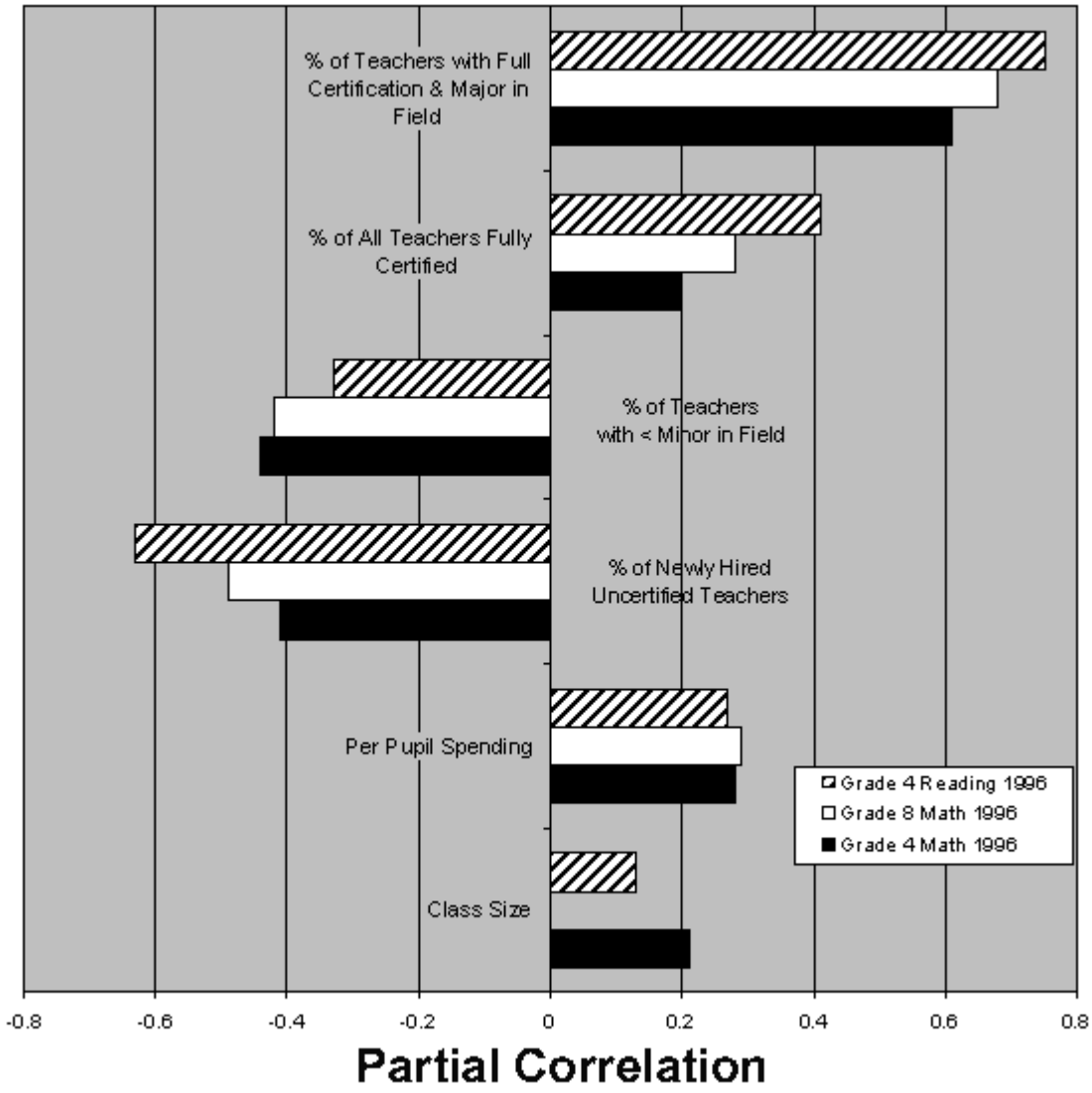

Figure 4. Partial Correlations (controlling for student poverty) between Selected Teacher Quality Variables and Student Achievement on the National Assessment of Educational Progress

Ordinary least squares regression analyses were performed to create the most parsimonious specification of a hyperplane of best fit with student achievement data. Because of the small sample size (n = 44 states participating in the state NAEP), the number of independent variables in each equation was minimized to preserve the necessary degrees of freedom (see Table 3). Variables were selected according to three criteria: to examine relationships often tested in other studies, to maximize explanatory power, and to avoid problems of multicollinearity. Teacher quality variables included the percentage of all teachers with full certification and a major in the field and the percentage of uncertified newly hired teachers, because these exhibit large influences on achievement, and the percentage of teachers with master's degrees, because this is a frequently examined teacher quality variable. Class size was also included because it is commonly found to influence achievement. Spending and salary variables were not included in the final estimations because they showed little relationship to student achievement in preliminary estimates. Because the percentage of minority students is

Analysis of Policy Relationships

Clearly, in any analysis such as this, the variables that can be measured are only proxies for the actual conditions or traits that may matter to student learning. In this case, a large number of variables associated with teacher quality appear to bear a significant relationship to student achievement. These include various ways of measuring state certification status (the proportions of teachers with full certification, less than full certification, and no certification) and disciplinary preparation (e.g., a major or minor in the field to be taught). Given the differences in licensing standards and teacher education programs across states, these proxies are fairly crude ones; nonetheless, they seem to indicate that teachers' knowledge, skills, and preparation matter for student achievement. The findings are similar to those of several other studies described earlier (Ferguson, 1991; Ferguson and Ladd, 1996; Fetler, 1999; Fuller, 1999; Strauss and Sawyer, 1986) in finding much stronger influences on student achievement of variables measuring teacher knowledge and skills than of variables like teacher experience, class sizes, or pupil-teacher ratios, which are generally found to have noticeable but smaller effects on student achievement where data are aggregated to the school or district levels.

The strength of the "well-qualified teacher" variable may be partly due to the fact that it is a proxy for both strong disciplinary knowledge (a major in the field taught) and substantial knowledge of education (full certification). If the two kinds of knowledge are interdependent as suggested in much of the literature, it makes sense that this variable would be more powerful than either subject matter knowledge or teaching knowledge alone. It is also possible that this variable captures other features of the state policy environment including general investments in, and commitment to, education, as well as aspects of the regulatory system for education, such as the extent to which standards are rigorous and the extent to which they are enforced. Recall that some states require teachers to acquire a subject matter major as well as extensive education training in human development and learning and in the methods of teaching in their field, while other states require much less extensive preparation in the content area as well as teaching and learning. In addition, some states are vigilant in enforcing their certification standards while others are not.

Teaching Standards and Other Policy Strategies

Finally, there may be unmeasured correlations between the extent to which states enact and enforce high standards for teachers and the extent to which they have enacted other policies that are supportive of public schools. Although it does not appear that teaching standards are strongly related to investments regarding class sizes or to overall education spending, it is possible that there are other factors influencing student achievement which generally co-exist with teacher quality and which were unmeasured in these estimates. Since most of the states which ranked among the highest-scoring on the NAEP examinations are strong local control states that have traditionally not exerted much control over school decision making, there are relatively few policy areas in which they have been active. Perhaps the relative lack of policy intervention is itself a support for student learning, leaving educators free of regulations that might force greater attention to procedures than learning. Another possibility is the influence of these states' small school and district sizes, a factor that has been identified in much research as contributing to student learning (for reviews, see Green & Stevens, 1988; Howley, 1989). In another analysis, Feistritzer (1993) has pointed out that most of the top-scoring states

on NAEP have very small average school sizes relative to national norms.

One area in which policies have not been positively correlated, however, is the extent to which states engaged in statewide student testing in the 1980s and the extent to which they enacted high standards for teachers. Among the 12 highest- scoring states in 8th grade mathematics in 1996 (10 of which had particularly high licensing standards in the form of subject matter and teaching coursework requirements), none had mandatory statewide testing programs in place during the 1980s or early 1990s. Only two of the top 12 states in 4th grade mathematics had statewide testing programs in place prior to 1995. By contrast, among the 12 lowest-scoring states (8 of which had particularly large rates of out-of-field and uncertified teachers), 10 had extensive student testing programs in place prior to 1990, some of which were associated with highly specified state curricula and an extensive menu of rewards and sanctions.

There are several possible interpretations of the almost inverse relationship between statewide testing policies and both teaching standards and student performance: It may be that states with low student performance and less qualified teachers were more likely to seek education improvements through student testing strategies and curriculum controls. It may also be that states have tended toward different theories of reform, with some investing more in testing and others in teaching. It is possible that regional differences in education investments and centralization happen to be correlated with policies regarding both testing and teacher investments (with Southern states that tend to score lowest investing heavily in curriculum and testing controls, while Northeastern and North Central states invest more in teacher education and less in curriculum controls).

The lack of apparent relationship between testing programs and student achievement might be because, without other investments to improve teaching and schooling, tests alone do not transform learning. Another possibility is that the kinds of basic skills tests and curricula enacted in many states during the 1980s were at odds with the NAEP assessments which increasingly seek to measure higher-order skills and performance abilities. It may be worth noting that most of the high-scoring and fast-gaining states discussed earlier instituted curriculum and testing reforms in the mid-1990s that were linked to the national student standards that guide NAEP and were much more performance-oriented than the basic skills tests that predominated in state assessment systems of the 1980s. While there is little evidence yet of the effects of these assessment programs on student learning, policy analysts may want to watch to see whether the types of tests matter for broad student outcomes as well as whether and how the supports that do or do not accompany testing programs (professional development, funding equalization, investments in additional supports for students ranging from early childhood education to special services of various kinds) make a difference.

Policies that May Influence Teachers' Qualifications

Another set of questions has to do with whether there are particular policy strategies used by states or districts that are associated with the preparation and hiring of better qualified teachers. The SASS data set and additional data collected directly from states allowed us to examine several policies in this regard.

Teacher education accreditation National data from the National Association of State Directors of Teacher Education and Certification and from the National Council for the Accreditation of Teacher Education provided the percentage of teacher education institutions that were accredited by NCATE. NCATE-accreditation might lead to higher overall standards for teachers because NCATE standards revisions in 1988 and 1993 required higher admissions standards, evidence of greater subject matter preparation, and

stronger rationales for the content of education coursework than those often emphasized by state approval systems.

Standard setting and enforcement mechanisms The state survey tracked the presence of a state professional standards board for teaching, analogous to the boards that govern other professions, which might enact and enforce higher standards. Since any policies for teacher education adopted by such a board would require several years to take broad effect, the enactment of a standards board prior to 1990 is the measure we used for examining influences on teacher qualifications in 1994.

District hiring standards SASS data provided the percentage of school districts in each state requiring each of the following as conditions for hiring: full state certification, graduation from an approved teacher education program, and a college major or minor in the field to be taught. There was wide variation across the states in the degree to which districts looked for evidence of these kinds of teacher qualifications as part of the hiring process.

Many more fine-grained variables, such as the content of licensing standards and the nature of teacher education programs, could not be tested with these data. Nonetheless, the results suggest some interesting associations. As shown in Table 4, the strongest predictor of the percentage of well-qualified teachers (that is, teachers with both a major and full certification in their field) is the percentage of teacher education institutions in a state that meet national accreditation standards through NCATE ($p < .05$).

Table 4
Relationship Between Professional Accreditation And Teacher Qualifications

Variable/ (Beta coefficient)	% of well-qualified teachers	% of well-qualified English teachers	% of well-qualified Math teachers	% of math teachers out-of-field	% of English teachers out-of-field
% of colleges NCATE accredited	.42*	.49**	.36*	-.37*	-.37*

*p<.05 **p<.01

The proportion of NCATE-accredited institutions is also significantly and negatively correlated with the proportion of English and mathematics teachers who are "out-of-field" (i.e., have less than a minor in the field they teach). This may be because institutions that are NCATE-accredited must demonstrate that their students have the opportunity to acquire a base of content knowledge deemed acceptable by the subject matter associations that review applications as well as pedagogical knowledge in their field. Thus, these institutions may, as a group, have less variability than others in establishing reasonably high standards for disciplinary knowledge as well as knowledge of how to teach the discipline. It may also be that states in which professional accreditation is more widespread also happen to have other policies or practices in effect that support the preparation and hiring of well-qualified teachers.

As shown in Table 5, the extent to which districts maintain rigorous hiring standards (i.e., the percentage of districts requiring full certification, graduation from an approved teacher education program, and a college major or minor in the field to be taught) is a highly significant predictor ($p < .001$) of the proportions of teachers who are uncertified. It is also a strong predictor of the proportions of new and veteran teachers who are fully certified. Since teachers' certification status is also related to state demographics, these variables were regressed against hiring standards along with student poverty, percent minority, and percent LEP students. The relationship between hiring standards and teacher certification status continues to be highly significant after controlling for student poverty, race, and language status.

Table 5
Correlations between Teacher Qualifications and District Hiring Standards (Pearson r)

District Hiring Standards (Percent of districts requiring full certification, graduation from an approved teacher education program, and a college major or minor in the field to be taught as a condition of hiring)	
% of new teachers who are fully certified	.28**
% of all teachers who are fully certified	.33**
% of newly hired teachers who are uncertified	-.51***
% of all teachers who are uncertified	-.66***

*p<.05 **p<.01 ***p<.001

Table 6
Relationship between Teacher Qualifications and District Hiring Standards
(Controlling for Student Poverty, Minority Status, and Language Status)

Variable/ Beta Weight/ (t value)	% all teachers fully certified	% new teachers fully certified	% all teachers uncertified	% new teachers uncertified
District hiring standards****	.393 (2.51)*	.339 (2.16)*	-.636 (-4.73)***	-.502 (-3.19)**
Professional Standards Board			-.173 (-1.20)	-.080 (-.48)
% students in poverty	-.148 (-.64)	-.063 (-.27)	.172 (.94)	-.108 (-.51)
% students LEP	.226 (1.23)	.374 (2.02)	.105 (.63)	.045 (.23)

% students minority	.125 (.58)	-.112 (-.43)	-.352 (-1.66)	-.105 (-.42)

*p<.05 **p<.01 ***p<.001

****Percent of districts requiring, as a condition of hiring, full certification, graduation from an approved teacher education program, and a college major or minor in the field to be taught

This suggests that enforcing standards is both a state and local job. In a quasi-profession like teaching, there is a complex interplay between the standards adopted by states and the ways in which local schools and districts manage their hiring processes, sometimes in accord with and sometimes in violation of state standards. A minority of states enforce their teacher licensing standards in the inviolable fashion with which standards for doctors, lawyers, architects, and other professionals are enforced. These other professions use professional standards boards established by each state as standard-setting and enforcement bodies. Depending on the degree of authority and autonomy used as defining characteristics, 12 to 18 states have established such boards for teaching.

As shown in Table 7, the presence of a professional standards board prior to 1990 proves to be significantly related to district hiring standards, a relationship that holds up after controlling for student characteristics. In addition, as Table 8 indicates, the presence of a standards board is significantly associated with the proportions of certified and uncertified teachers. This relationship may work through the influence such a board exerts over district decisions about hiring qualified personnel, as suggested above. Districts often hire unqualified teachers even though fully prepared teachers are available if state agencies do not prevent them from doing so. This can occur as a function of cumbersome hiring procedures, patronage, lack of recruitment effort or incentives, or efforts to reduce salary costs (NCTAF, 1996). Depending upon how they are structured, some standards boards may have more authority and/or more commitment to prevent the hiring of unqualified teachers than some state agencies do. In agency interviews, for example, a staff member of a highly effective state standards board described how the board examines the candidate qualifications as well as the district's advertising, selection, and hiring practices and applicant pool in any case where a district requests permission to hire staff on an emergency or temporary license. Very few requests for hiring of unqualified personnel are ultimately granted, and district hiring practices are often revised and improved in the process of the review. In other states, agency officials described routine, blanket approvals of district requests for emergency hiring even in situations where districts had just laid off large numbers of qualified teachers or had qualified applicants in the applicant pool. These officials generally felt they did not have the resources or the authority to investigate or stem practices they felt were illegal and widespread.

Table 7
Correlations (Pearson r) of Presence of a Professional Standards Board with District Hiring Standards and Teacher Qualifications

% of districts requiring graduation from an approved teacher education program	.25*

% of districts requiring a college major or minor in the field to be taught	.23*
% of districts requiring full certification, graduation from an approved program, and a college major or minor	.30**
% uncertified teachers	-.27**
% fully certified teachers	.21*
% fully certified new teachers	.21*
# of weeks required for student teaching	.25*

*p<.05 **p<.01

Table 8
Relationship between Professional Standards Board Presence and District Hiring Standards

	District hiring standards
Professional Standards Board	.411 (2.49)**
% students in poverty	.132 (.58)
% LEP students	-.429 (-2.20)*
% minority students	.067 (.26)

*p<.05 **p<.01

These relationships between the presence of standards boards and teacher education or hiring practices, although statistically significant, are quite modest (correlations in the .2 to .3 range), suggesting that many other variables are at play here as well. It is certainly true that some states enact and enforce high standards for teaching without the presence of standards boards, while some standards boards do not pursue their mission with the same vigor as others. Where they exist, however, such bodies often appear to bring greater consistency of effort and attention to the issues of preparation and qualifications.

Conclusions and Implications

This analysis triangulates data from surveys of state policies, case study analyses of state policymaking, and quantitative examination of the distribution of state achievement scores and resources, taking student characteristics into account. Some findings are

particularly noteworthy. First, while student demographic characteristics are strongly related to student outcomes at the state level, they are less influential in predicting achievement levels than variables assessing the quality of the teaching force. Second, when aggregated at the state level, teacher quality variables appear to be more strongly related to student achievement than class sizes, overall spending levels, teacher salaries (at least when unadjusted for cost of living differentials), or such factors as the statewide proportion of staff who are teachers.

Among variables assessing teacher "quality," the percentage of teachers with full certification and a major in the field is a more powerful predictor of student achievement than teachers' education levels (e.g., master's degrees). This finding concurs with those of other studies cited earlier. It is not surprising that masters degrees would be relatively weaker measures of teacher knowledge, given the wide range of content they can include, ranging from specialist degrees in reading or special education that are directly related to teaching to fields like administration and others that have little to do with teaching. Other measures of certification status (e.g., the percent of teachers uncertified, the percent with full certification) are also strong correlates of student achievement. Finally, certain policy strategies associated with standard-setting at the state and local level--NCATE-accreditation of teacher education institutions, district hiring standards, and, to a lesser extent, state professional standards boards--appear to be related to teacher qualifications in the field.

While the triangulation of data from several sources lends some confidence to these findings, they should be viewed with caution. Like all studies that draw inferences from broad state trends and correlational data, there are many variables in play at any given time and many possible explanations for any phenomenon observed. While this article presents a range of competing explanations for student achievement trends (student background, curriculum and testing policies, school funding and equalization, school and class sizes), it could not fully test all of these explanations. This remains for other researchers to pursue. In addition, other data and other methodologies could shed further light on these questions. Adding information about parent education levels might make a difference in the measurement of student background; adding data about school and district size (from the Common Core of Data) and curriculum and testing approaches (from the NAEP background surveys) might shed greater light on school factors that matter; and adjusting salary and spending data for cost of living differentials might allow a better evaluation of fiscal influences.

By including estimates of the proportions of staff who are underqualified (and who tend to cluster in less advantaged schools and districts), this study's estimates tapped some of the local variability in resources made available to children. However, because state data on average class sizes and other school resources ignore wide variations in teaching and learning conditions that may be very important at the district, school, and classroom levels, these estimates cannot fully capture the effects of such variables. Average class sizes, for example, vary relatively little across states but vary substantially within states and districts. Thus, effects of this variable are much more likely to be perceived with more disaggregated data. By merging district, school, and teacher files, the SASS data can allow for the use of Hierarchical Linear Modeling techniques, which would be a useful tool for further exploring relationships between teaching and schooling variables at the school, district, and state levels.

Nonetheless, the findings of this study, in conjunction with a number of other studies in recent years, suggest that states interested in improving student achievement may be well-advised to attend, at least in part, to the preparation and qualifications of the teachers they hire and retain in the profession. It stands to reason that student learning

should be enhanced by the efforts of teachers who are more knowledgeable in their field and are skillful at teaching it to others. Substantial evidence from prior reform efforts indicates that changes in course taking, curriculum content, testing, or textbooks make little difference if teachers do not know how to use these tools well and how to diagnose their students' learning needs (for a review, see Darling-Hammond, 1997b).

Like other studies cited earlier, this research indicates that the effects of well-prepared teachers on student achievement can be stronger than the influences of student background factors, such as poverty, language background, and minority status. And while smaller class sizes appear to contribute to student learning, particularly in fields like elementary reading, the gains occasioned by smaller classes are most likely to be realized, as they were in the Tennessee experiment, when they are accompanied by the hiring of well-qualified teachers. The large- scale hiring of unqualified teachers, as was the case in California's recent class size reduction initiative, would likely offset any achievement gains that could be realized by smaller class sizes.

Another implication of this study is that states may impact the qualifications of the teachers through policies that influence the hiring standards of school districts (e.g., incentives and sanctions from the state level that encourage the hiring of well-qualified individuals), the accreditation of teacher education institutions (e.g., encouragement or requirements for the use of NCATE standards or others of equivalent rigor), and the bodies that establish and enforce teaching standards (e.g. establishment of professional standards boards or assurance of adequate capacity and authority for state agencies to uphold high standards for teaching).

Although this study used fairly crude measures of teacher knowledge and skills such as certification status, college major, and master's degrees, policymakers should be aware that there are much more fine-grained distinctions to be made among types of state certification standards, teacher education programs, professional development offerings, and education requirements that make a difference to the teachers' abilities and their students' outcomes. Reforms underway to create more thoughtful licensing systems, more productive teacher education programs, and more effective professional development strategies are producing evidence of the stronger effects on teaching and learning of approaches that strengthen teachers' abilities to teach diverse learners with a keen diagnostic eye and a wide repertoire of strategies supporting mastery of challenging content (for a review, see NCTAF, 1996; Darling-Hammond 1997a). Over the next decade, federal, state, and local policymakers interested in helping students meet higher learning standards may want to consider how investments in teacher quality, along with other reforms, can assist them in achieving their goals.

Notes

1. This research was funded in part by the Office of Educational Research and Improvement (OERI) of the U.S. Department of Education through the Center for the Study of Teaching and Policy, which is housed at the University of Washington and includes Stanford University, Teachers College, Columbia University, and the University of Michigan. The research was initiated while the author was a fellow at the Center for Advanced Study in the Behavioral Sciences with the support of the Spencer Foundation. The views represented in this article are those of the author alone, and do not represent those of any sponsor.
2. National Center for Education Statistics, *Schools and Staffing Survey, 1993-94: State by State Data*, Washington, DC: U.S. Department of Education, 1996, Table 3.5. Additional tabulations performed by the National Commission on Teaching

and America's Future.

3. The INTASC standards, developed by a consortium of more than 30 states and professional associations under the auspices of the Council of Chief State School Officers, are based on knowledge of effective learning and teaching and on the student learning standards developed by professional associations such as the National Council of Teachers of Mathematics. The INTASC standards for beginning teacher licensing are compatible with the more advanced standards of the National Board for Professional Teaching Standards, which define and assess accomplished teaching among veteran teachers.

References

Adams, R.D., S. Hutchinson, and C. Martray (1980). A developmental study of teacher concerns across time. Paper presented at the American Educational Research Associaton Annual Meeting. Boston, Mass.

Andrew, M. & Schwab, R.L. (1995). Has reform in teacher education influenced teacher performance? An outcome assessment of graduates of eleven teacher education programs. *Action in Teacher Education, 17*, 43-53.

Andrews, J.W., Blackmon, C.R., and Mackey, J.A. (1980). Preservice performance and the National Teacher Examinations. *Phi Delta Kappan, 61*(5), pp. 358-359.

Armour-Thomas, E., Clay, C., Domanico, R., Bruno, K., & Allen, B. (1989). An outlier study of elementary and middle schools in New York City: Final report. NY: New York City Board of Education.

Ashton, P. & Crocker, L. (1987, May-June). Systematic study of planned variations: The essential focus of teacher education reform. *Journal of Teacher Education, 38*, 2-8.

Ayers, J.B., and Qualls, G.S. (Nov/Dec 1979). Concurrent and predictive validity of the National Teacher Examinations. *Journal of Educational Research, 73* (2), pp.86-92.

Barnes, S., Salmon, J. and Wale, W. (1989), "Alternative Teacher Certification in Texas." Presented at the annual meeting of the American Educational Research Association, March. ERIC Document No. 307316.

Begle, E.G. (1979). Critical Variables in Mathematics Education. Washington, D.C.: Mathematical Association of American and National Council of Teachers of Mathematics.

Begle, E.G. and Geeslin, W. (1972). Teacher effectiveness in mathematics instruction. National Longitudinal Study of Mathematical Abilities Reports No. 28. Washington, D.C. Mathematical Association of America and National Council of Teachers of Mathematics.

Bents, Mary, and Richard Bents (1990). Perceptions of Good Teaching Among Novice, Advanced Beginner and Expert Teachers. Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

Berliner, D.C. & Biddle, B.J. (1995). *The manufactured crisis: Myth, fraud, and the*

attack on America's public schools. Reading, MA: Addison-Wesley.

Berliner, D.C. & Tikunoff, W. J. (1976). The California Beginning Teacher Study. *Journal of Teacher Education, 27* , 24-30.

Bowles, S., and Levin, H.M. (1968). The determinants of scholastic achievement—An appraisal of some recent evidence. *Journal of Human Resources, 3*, 3-24.

Brown, C.A., Smith, M.S., and Stein, M.K. (1995). Linking teacher support to enhanced classroom instruction. Paper presented at the annual meeting of the American Educational Research Association. New York, New York.

Byrne, C.J. (1983). Teacher knowledge and teacher effectiveness: A literature review, theoretical analysis and discussion of research strategy. Paper presented at the meeting of the Northwestern Educational Research Association, Ellenville, NY.

Campbell, J.R., Donahue, P.L., Reese, C.M., Phillips, G.W. (1996). NAEP 1994 reading report card for the nation and the states. Washington, D.C.: U.S. Department of Education.

Carroll, J.B. (1975). *The teaching of French as a foreign language in eight countries.* NY: John Wiley and Sons.

Carter, K., and Doyle, W. (1987). Teachers' knowledge structures and comprehension processes. Pp. 147-160 in J. Calderhead (Ed.), *Exploring Teacher Thinking*. London: Cassell.

Connecticut State Department of Education Division of Research, Evaluation, and Assessment (1991). Research Bulletin, School Year 1990-91, No. 1 (Hartford, CT: Bureau of Research and Teacher Assessment).

Cooper, E. & Sherk, J. (1989). Addressing urban school reform: Issues and alliances. *Journal of Negro Education, 58*, 3, 315-331.

Cohen, D.K. and Hill, H. (1997). Instructional Policy and Classroom Performance: The Mathematics Reform in California. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

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

Council on School Performance (1997). Teachers with advanced degrees advance student learning. Atlanta: Council for School Performance, Georgia State University.

Darling-Hammond, L. (1992). Teaching and knowledge: Policy issues posed by alternative certification for teachers. *Peabody Journal of Education, 67*, 3, 123-154.

Darling-Hammond, L. (1997a). Doing What matters most: Investing in quality teaching. NY: National Commission on Teaching and America's Future.

Darling-Hammond, L. (1997b). *The right to learn: A blueprint for creating schools that*

work. San Francisco: Jossey Bass.

Darling-Hammond, L.; Hudson, L.; and Sheila Kirby (1989). Redesigning Teacher Education: Opening the Door for New Recruits to Science and Mathematics Teaching. Santa Monica:The RAND Corporation.

Darling-Hammond, L., Wise, A.E. & Klein, S. (1995). *A license to teach: Building a profession for 21st century schools*. Boulder: Westview Press.

Darling-Hammond, L., Wise, A.E. and Pease S.R. (1983). Teacher evaluation in the organizational context: a review of the literature. *Review of educational research, 53*, 285-237.

Denton, J.J., and L.J. Lacina (1984). Quantity of professional education coursework linked with process measures of student teaching. *Teacher Education and Practice*, 39-64.

Denton, J.J., and Peters, W.H. (1988). Program assessment report: Curriculum evaluation of a non-traditional program for certifying teachers. Texas A &M University, College Station, TX.

Doyle, W. (1985). Recent research on classroom management: Implications for teacher preparation. *Journal of Teacher Education, 36*, 3, 31-35.

Doyle, W. (1986). Content representation in teachers' definitions of academic work. *Journal of Curriculum Studies, 18*, 365- 379.

Druva, C.A., & Anderson, R.D. (1983). Science teacher characteristics by teacher behavior and by student outcome: A meta-analysis of research. *Journal of Research in Science Teaching, 20*,5, 467-479.

Erekson, T.L. and Barr, L. (1985). Alternative credentialing: Lessons from vocational education. *Journal of Teacher Education, 36*, 3, 16-19.

Evertson, C., Hawley, W., & Zlotnik, M. (1985). Making a difference in educational quality through teacher education. *Journal of Teacher Education, 36*,3, 2-12.

Feiman-Nemser, Sharon, and Michelle B. Parker (1990). Making Subject Matter Part of the Conversation or Helping Beginning Teachers Learn to Teach. East Lansing, MI: National Center for Research on Teacher Education.

Feistritzer, C.E. (1993). Report Card on American Education: A State-by-State Analysis, 1972-73 to 1992-93. Washington, D.C.: National Center on Education Information.

Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives, 7*, 9 [Entire issue]. (Available online at http://epaa.asu.edu/epaa/v7n9.html)

Ferguson, R.F. (1991, Summer). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation, 28*,2, 465-498.

Ferguson, R.F. and Ladd, H.F. (1996). How and Why Money Matters: An Analysis of

Alabama Schools. Pp. 265-298 in Helen Ladd (Ed.) *Holding Schools Accountable*. Washington, D.C.: Brookings Institution.

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

Fuller, E. J. (1999). Does teacher certification matter? A comparison of TAAS performance in 1997 between schools with low and high percentages of certified teachers. Austin: Charles A. Dana Center, University of Texas at Austin.

Glass, G.V; Cahen, L.S.; Smith, M.L.; and Filby, N.N. (1982). *School class size: Research and Policy.* Beverly Hills, CA: SAGE Publications.

Glassberg, S. (1980). A view of the beginning teacher from a developmental perspective. Paper presented at the American Educational Research Association Annual Meeting. Boston, MA.

Goebel, Stephen D., Karl Ronacher, and Kathryn S. Sanchez (1989), An Evaluation of HISD's Alternative Certification Program of the Academic Year: 1988-1989. Houston: Houston Independent School District Department of Research and Evaluation. ERIC Document No. 322103.

Goldhaber, M. and Brewer, D. (1999). Does teacher certification matter? High school certification status and student achievement. Unpublished manuscript.

Gomez, D.L., and Grobe, R.P. (1990). Three Years of Alternative Certification in Dallas: Where are We? Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

Good, T.L. (1983). Recent classroom research: Implications for teacher education. In D.C. Smith (ed.), *Essential knowledge for beginning educators*. Washington, D.C.: American Association of Colleges for Teacher Education.

Good, T.L. and J. E. Brophy. (1986). *Educational Psychology, 3rd Edition*. NY: Longman.

Grady, M.P., Collins, P. & Grady, E.L. (1991). Teach for America 1991 Summer Institute Evaluation Report. Unpublished manuscript.

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

Grissmer, D. and Flanagan, A. (1998). Exploring Rapid Achievement Gains in North Carolina and Texas. Washington, D.C.: National Education Goals Panel.

Grossman, P. L. (1989). Learning to teach without teacher education. *Teachers College Record, 91*, 2, 191-208.

Guyton, E. and Farokhi, E. (1987). Relationships among academic performance, basic skills, subject matter knowledge and teaching skills of teacher education graduates. *Journal of Teacher Education, 38,* (Sept-Oct.), 37-42.

Haberman, M. (1995, June). Selecting 'star' teachers for children. *Phi Delta Kappan, 76*, 777-781.

Haney, W. (1999). Supplementary report on Texas Assessment of Academic Skills Exit Test (TAAS-X). Boston: Center for the Study of Testing, Evaluation, and Educational Policy.

Haney, W., Madaus, G., & Kreitzer, A. (1987). Charms talismanic: testing teachers for the improvement of american education. Pp. 169-238 in E.Z. Rothkopf (Ed.) *Review of Research in Education*, Vol. 14. Washington, D.C.: American Educational Research Association.

Hansen, J. B. (1988). The Relationship of Skills and Classroom Climate of Trained and Untrained Teachers of Gifted Students. Unpublished doctoral dissertation, Purdue University.

Hanushek, E.A. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *The American Economic Review, 61*(2), 280-288.

Hanushek, E.A., Rivkin, S.G., and Taylor, L.L. (1995). Aggregation bias and the estimated effects of school resources. Working paper 397. University of Rochester, Center for Economic Research.

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

Hellfritzch, A.G. (1945). A factor analysis of teacher abilities. *Journal of Experimental Education, 14*, 166- 169.

Hice, J.E.L. (1970). The Relationship between Teacher Characteristics and First-Grade Achievement. *Dissertation Abstracts International, 25*, 1, 190.

Johanson, G.A. and Gips, C.J. (1992). The hiring preferences of secondary school principals. *The High School Journal, 76*, (Oct./Nov.), 1-16.

Jordan, H.R., Mendro, R.L., & Weersinghe, D. (1997). Teacher effects on longitudinal student achievement: A preliminary report on research on teacher effectiveness. Paper presented at the National Evaluation Institute, Indianapolis, IN. Kalamazoo, MI: CREATE, Western Michigan University.

Kentucky Institute for Education Research (1997). The preparation of teachers for Kentucky Schools: A survey of new teachers. Frankfort, KY: Kentucky Institute for Education Research

Klitgaard, R.E. and Hall, G.R. (1974). Are there unusually effective schools? *Journal of Human Resources, 10*, 3, 90-106.

Knoblock, G.A. (1986). Continuing professional education for teachers and its relationship to teacher effectiveness. Unpublished dissertation. Michigan State University. *Dissertation Abstracts International, 46*, 2, 3325A, University microfilms no. AAC8529729.

Kurtz, M. (1999). Schools, teens battle barrier of 9th grade. *American-Statesman*, May 23, 1999.

LaDuke, D.V. (1945). The measurement of teaching ability. *Journal of Experimental Education, 14*, 75-100.

Lenk, Harriet Anne (1989). A Case Study: The Induction of Two Alternate Route Social Studies Teachers. Unpublished doctoral dissertation. Teachers College, Columbia University.

LuPone, L.J. (1961). A comparison of provisionally certified and permanently certified elementary school teachers in selected school districts in New York State. *Journal of Educational Research, 55*, 1, 53-63.

Lutz, Frank W., and Jerry B. Hutton (1989). Alternative teacher certification: Its policy implications for classroom and personnel practice. *Educational Evaluation and Policy Analysis, 11*, 3, 237-254.

McNeil, J.D. (1974). Who gets better results with young children -experienced teachers or novices? *Elementary School Journal, 74*, 447-451.

Mitchell, N. (1987). Interim Evaluation Report of the Alternative Certification Program (REA87-027-2). Dallas, TX: DISD Department of Planning, Evaluation, and Testing.

Monk, D. H. (1994). Subject matter preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*, 2, 125-145.

Monk, D. H. and King, J.A. (1994). Multilevel teacher resource effects in pupil performance in secondary mathematics and science: The case of teacher subject matter preparation. Pp. 29- 58 in R.G. Ehrenberg (Ed.), *Choices and consequences: Contemporary policy issues in education*. Ithaca, NY: ILR Press.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children, 5*, 2, 113- 127.

Murnane, R.J. (1985, June). Do effective teachers have common characteristics: Interpreting the quantitative research evidence. Paper presented at the National Research Council Conference on Teacher Quality in Science and Mathematics, Washington, D.C.

Murnane, R.J., and 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, 1fR,* 4, 691-693.

National Center for Education Statistics (1994). Data compendium for the NAEP 1992 reading assessment of the nation and the states: 1992 NAEP trial state assessment. Washington, D.C.: U.S. Department of Education.

National Center for Education Statistics (n.d.) NAEP 1992, 1994 national reading assessments – Data almanac – Grade 4: Teacher questionnaire weighted percentages and composite proficiency means (Public school). (Available online at http:// www.Nces.ed.gov/nationsreportcard/y25alm/almanac.shtml)

National Association of State Directors of Teacher Education and Certification (NASDTEC)(1997). Manual on certification and preparation of educational personnel in the United States and Canada, 1997-98.

National Center for Education Statistics (NCES) (1995). Digest of Education Statistics, 1995. Washington, D.C.: U.S. Department of Education.

National Commission on Teaching and America's Future (NCTAF) (1996). What Matters Most: Teaching for America's Future. New York: Author.

Natriello, G., Zumwalt, K., Hansen, A., and Frisch, A. (1990). Characteristics of Entering Teachers in New Jersey. Revised version of a paper presented at the 1988 Annual Meeting of the American Educational Research Association.

Olsen, D. G. (1985). The quality of prospective teachers: Education vs. noneducation graduates, *Journal of Teacher Education, 36*, 5, 56-59.

Penick, J.E., and Yager, R.E. (1983). The search for excellence in science education. *Phi Delta Kappan, 64*, 8, 621-623.

Perkes, V.A. (1967-1968). Junior high school science teacher preparation, teaching behavior, and student achievement. *Journal of Research in Science Teaching, 6*, 4, 121-126.

Pflaum, S.W. and Abramson, T. (1990, March). Teacher assignment, hiring, and preparation: Minority teachers in New York City. *The Urban Review, 22*, 17-31.

Quirk, T.J., Witten, B.J., and Weinberg, S.F. (1973). Review of studies of concurrent and predictive validity of the National Teacher Examinations. *Review of Educational Research, 43,* 89-114.

Reese, C.M., Miller, K.E., Mazzeo, J., Dossey, J.A. (1997). NAEP 1996 Mathematics Report Card for the Nation and the States. Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

Rosenholtz, S. J. (1986). The organizational context of teaching. In *Learning to Teach*. University of Illinois at Champaign-Urbana.

Rosenshine, B. and Furst, N.F. (1973). The use of direct observation to study teaching. In R.M.W. Travers (ed.), *Handbook of Research on Teaching (2nd Edition)*. Chicago: Rand McNally.

Rostker, L.E. (1945). The measurement of teaching ability. *Journal of Experimental Education, 14*, 5-51.

Rottenberg, Claire J. and David C. Berliner (1990). Expert and Novice Teachers' Conceptions of Common Classroom Activities. Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

Roupp, R.; Travers, J.; Glantz, F.; and Coelen, C. (1979). Children at the Center: Summary Findings and their Implications. Cambridge, MA: Abt Associates.

Sanders, S.L., Skonie-Hardin, S.D., and Phelps, W.H. (1994, November). The effects of teacher educational attainment on student educational attainment in four regions of Virginia: Implications for administrators. Paper presented at the Annual Meeting of the Mid-South Educational Research Association.

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

Schalock, D. (1979). Research on teacher selection. In D.C. Berliner (Ed.), *Review of research in education, Vol. 7,* Washington, D.C.: American Educational Research Association.

Skinner, W.A. (1947). An investigation of factors useful in predicting teaching ability. University of Manchester. Master of Education thesis.

Skipper, C. E. and Quantz, R. (1987). Changes in educational attitudes of education and arts and science students during four years of college, *Journal of Teacher Education, 38*, May- June, 39-44.

Soar, R.S., Medley, D.M., and Coker, H. (1983). Teacher evaluation: A critique of currently used methods. *Phi Delta Kappan, 65*, 4, 239-246.

Stoddart, T. (1992). An alternate route to teacher certification: Preliminary findings from the Los Angeles Unified School District Intern Program. *Peabody Journal of Education,67*, 3.

Strauss, R. P. and Sawyer, E.A. (1986). Some New Evidence on Teacher and Student Competencies. *Economics of Education Review, 5*, 1, 41-48.

Summers, A.A., and Wolfe, B.L. (1975, February). Which School Resources Help Learning? Efficiency and Equality in Philadelphia Public Schools. Philadelphia, PA: ED 102 716

Taylor, J.K. and R. Dale (1971). A Survey of Teachers in the First Year of Service. Bristol: University of Bristol, Institute of Education.

Texas Education Agency (1993). Teach for America Visiting Team Report. Austin: Texas State Board of Education Meeting Minutes, Appendix B.

Vernon, P.E. (1965). Personality Factors in Teacher Trainee Selection. *British Journal of Educational Psychology, 35,* 140-149.

Walberg, H.J., Waxman, H.C. (1983). Teaching, learning, and the management of instruction. In D.C. Smith (Ed.), Essential knowledge for beginning educators. Washington, DC: American Association of Colleges for Teacher Education and ERIC Clearinghouse on Teacher Education.

Wiley, D. & Yoon, B. (1995). Teacher reports of opportunity to learn: Analyses of the 1993 California Learning Assessment System. *Educational Evaluation and Policy Analysis, 17*, 3, 355-370.

Wise, A.E.; Darling-Hammond, L.; and Berry, B. (1987). Effective Teacher Selection, From Recruitment to Retention. Santa Monica, CA: RAND Corporation.

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

About the Author

Linda Darling-Hammond
School of Education
Stanford University

Email: ldh@leland.stanford.edu

Linda Darling-Hammond is Charles E. Ducommun Professor of Education at Stanford University and executive director of the National Commission on Teaching and America's Future. Her research, policy, and teaching focus on teacher education and teaching quality, school restructuring, and educational equity. Among other writings, she is author of *The Right to Learn,* which received the Outstanding Book Award from the American Educational Research Association in 1998.

EPAA Spanish Language Editorial Board