

# Education Policy Analysis Archives

Volume 10 Number 37

September 6, 2002

ISSN 1068-2341

---

A peer-reviewed scholarly journal

**Editor: Gene V Glass**

College of Education

Arizona State University

Copyright 2002, the **EDUCATION POLICY ANALYSIS ARCHIVES**.

Permission is hereby granted to copy any article if **EPAA** is credited and copies are not sold. **EPAA** is a project of the [Education Policy Studies Laboratory](#).

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*.

---

## **The Effectiveness of "Teach for America" and Other Under-certified Teachers on Student Academic Achievement: A Case of Harmful Public Policy<sup>1</sup>**

**Ildiko Laczko-Kerr**

**Arizona Department of Education**

**David C. Berliner**

**Arizona State University**

Citation: Laczko-Kerr, I., & Berliner, D.C.. (2002, September 6). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy," *Education Policy Analysis Archives*, 10(37). Retrieved [date] from <http://epaa.asu.edu/epaa/v10n37/>.

### **Abstract**

The academic achievements of students taught by under-certified primary school teachers were compared to the academic achievements of students taught by regularly certified primary school teachers. This sample of under-certified teachers included three types of under-qualified personnel: emergency, temporary and provisionally

certified teachers. One subset of these under-certified teachers was from the national program "Teach For America (TFA)." Recent college graduates are placed by TFA where other under-qualified under-certified teachers are often called upon to work, namely, low-income urban and rural school districts. Certified teachers in this study were from accredited universities and all met state requirements for receiving the regular initial certificate to teach. Recently hired under-certified and certified teachers (N=293) from five low-income school districts were matched on a number of variables, resulting in 109 pairs of teachers whose students all took the mandated state achievement test. Results indicate 1) that students of TFA teachers did not perform significantly different from students of other under-certified teachers, and 2) that students of certified teachers out-performed students of teachers who were under-certified. This was true on all three subtests of the SAT 9—reading, mathematics and language arts. Effect sizes favoring the students of certified teachers were substantial. In reading, mathematics, and language, the students of certified teachers outperformed students of under-certified teachers, including the students of the TFA teachers, by about 2 months on a grade equivalent scale. Students of under-certified teachers make about 20% less academic growth per year than do students of teachers with regular certification. Traditional programs of teacher preparation apparently result in positive effects on the academic achievement of low-income primary school children. Present policies allowing under-certified teachers, including those from the TFA program, to work with our most difficult to teach children appear harmful. Such policies increase differences in achievement between the performance of poor children, often immigrant and minority children, and those children who are more advantaged.

There has been growing interest in "teacher quality," a catch phrase for a host of teacher characteristics, including a teacher's content knowledge, classroom behavior, academic ability, advanced degree work, salary, and teacher education experiences. Among the many characteristics under investigation as an indicator of teacher quality has been teacher certification. This study examines the effects of different kinds of teacher certification on student achievement. Reviews of this issue may be found in Darling-Hammond, 1999 and 2002; Evertson, Hawley & Zlotnik, 1985; and Ashton, Crocker, & Olejnik, 1986.

In Arizona, a state with a strong commitment to standards based reform, policies were needed to ensure that quality teachers would be available for students to meet the new and more rigorous mandated standards. Thus the Arizona Educator Proficiency Assessment (AEPA) was developed as a tool in the state certification process to ensure the quality of new teachers. One part of the test purports to measure teachers' professional knowledge, including pedagogy, teaching methods, and educational theory. The second part of the test covers content knowledge, either elementary content, or for secondary teachers, a subject specific content area. A passing score on the test, clearance by the police of any criminal record, and an accredited university's recommendation that a person is prepared to work as a classroom teacher earns a regular certificate to teach from the State.

Arizona's efforts are part of a national movement to improve the quality of teachers through

assessments like the AEPA ("Quality Counts," 2000; Higher Education Reauthorization Act, 1998). But not every district in the state or nation can find regularly certified teachers, giving rise to other policies that appear to work against the goal of increased teacher quality. For example, in Arizona and elsewhere, attempts at improving the quality of the teaching force seem contradicted by the continuing practice of issuing emergency certification (see "Quality Counts," 2000; Olson, 2000). Critics of hiring uncertified teachers ask whether complex, standards—based reforms can be enacted with teachers who are, to varying degrees, untrained. Supporters of hiring uncertified teachers claim that the advantages of traditional teacher education programs are unproven, and some question, as well, whether such training is even necessary. Stated in its simplest form as a research problem the question is: "Do students taught by teachers with emergency certification learn as much or achieve as well as students who are taught by regularly certified teachers?" An answer to this simple question would inform us whether policies designed to improve teacher quality are being undermined by the simultaneous adoption of policies that allow the use of uncertified teachers.

The dilemma associated with using uncertified teachers is not limited to Arizona where, currently, one out of six teachers are estimated to be uncertified (Go, 2002). For example, the Chicago "Sun Times" (Rossi & Grossman, 2002) reports an audit by the Chicago Board of Education showing that 22 percent of teachers in the system's 81 probationary schools—those with the greatest academic needs and the lowest test scores—were not fully qualified to teach. These were teachers missing what the state calls "initial" or "standard" certificates. Other teachers were found with certificates, but they were teaching subjects they were not certified to teach. In all, 900 teachers, about one of every five in Chicago's worst-performing public schools appeared unqualified to teach during the school year 2001-2002. New York State appears to be no different. Lankford, Loeb and Wycoff (2002) report that in a recent school year, in some New York schools, less than half the teachers held certification for the courses they taught. These schools were invariably urban and serving the poor.

With the passage of the *No Child Left Behind Act of 2001* (HR1), federal law will require schools to have a "highly qualified" teacher in every classroom by 2005-2006. Thus research on the effectiveness of uncertified and certified teachers takes on added significance as the designation of teachers as uncertified and certified becomes entwined with the evaluations of teachers so that the highly qualified can be distinguished from those less qualified.

## **Related Research**

The review that follows explores the difference between licensure and certification, reasons for teacher certification, the role of certification in the professionalization of the teaching workforce, on-the-job training for teachers, and the assessment of beginning teachers' competency. After those topics are discussed briefly, research that bears on a broad range of certification issues and teacher effectiveness is discussed in somewhat more detail. These areas are all highly contested and interpretations of this literature are, more than usual, intertwined with the ideology of the researchers (Cochran-Smith & Fries, 2001).

### **Licensure and Certification**

Teacher certification, at its core, is based on the need to ensure that every public school teacher has had rigorous screening and training and been judged qualified to teach. Certification is designed to protect the public from harm. But there is a difference between certification and licensure. Lawyers, cosmetologists, and physicians represent a few of the many professions that

require a practitioner to hold a license in order practice their profession. The goal of licensing tests is to set a minimum level of competency. Professions that require licensure make it illegal for someone without a license to practice that occupation (Pyburn, 1990). A person without a law or cosmetology license would be committing a crime if caught practicing law or working in a hair-dressing salon. No such legal protection is afforded the public when it comes to education. Teachers without certification are simply not allowed to use the title of "certified teacher" but there are no legal impediments for teaching without certification. This difference between certification and licensure allows states to issue emergency certificates but not emergency licenses. Issuance of these emergency certificates produces considerable moral difficulty. A newspaper report on Chicago public schools makes this case dramatically (Rossi & Grossman, 2002). The journalists document that at Howland elementary school, one of Chicago's poorest:

55 percent of teachers were "not fully certified" to teach all of their students.... That included four of six teachers in high-stakes classrooms, where kids must repeat a grade if they don't do well on annual tests. All four held substitute certificates, although two were in teacher preparation programs. Assigning uncertified substitutes to low-scoring kids who face high-stakes tests "should be illegal," said Kati Haycock, head of the Education Trust, a Washington, D.C., research and advocacy group. "That's educational malpractice." However, in Chicago, no policy governs who can teach such students.

The "Philadelphia Inquirer" wondered what the city of Philadelphia was going to do about the same problem (Hundreds of teachers, 2002). President George W. Bush had just asked America to ensure that there was a highly qualified teacher in every classroom. But Philadelphia has 30,000 students in classrooms where teachers are uncertified and the city cannot figure out how to solve that problem. President Bush did not mention that while Governor of Texas, during the 1996-97 school year, he allowed 760,000 of the state's 3.8 million students to be taught by uncertified teachers. Nor did he note that the students with uncertified teachers were found not to do as well on the state achievement tests as did students in the classes of regularly certified teachers (Students of certified, 1999). President George W. Bush has now passed on to the nation the problem that Governor George W. Bush could not solve.

## **Reasons for Certification**

Those who defend the process of teacher certification claim it is a necessary component in the development and maintenance of the profession of teaching, as well as the means by which the state can ensure the quality of those who enter the profession. Wise (1994a) notes, however, that there are two methods of controlling entry into the profession, professional control and popular control.

Professional control allows the teaching profession to monitor who becomes a teacher. By specifying standards for certification and through various political mechanisms, the profession controls the quality of teachers who enter the profession. When professional control is present we often see teacher input in the design of teacher certification tests. On the other hand, popular control allows public demand to control who is placed in classrooms, with much less concern for their qualifications. Emergency certificates to teach during times of "shortages" are an example of popular control. Wise (1994b) advocates professional control as the primary means to allow the promotion of teaching within the economic sector. He believes that without

certification teaching becomes a trade rather than a profession.

Among other characteristics, a profession is also defined as possessing a distinct body of knowledge and having control of the education and licensing of its members (Pratte & Rury, 1991; Burbules & Densmore, 1991). Labaree (1992) describes professionalization as the ability to demonstrate formal knowledge and to have autonomy in the work place. He explains that any occupational group:

...must establish that it has mastery of a formal body of knowledge that is not accessible to the layperson and that gives it special competence in carrying out a particular form of work. In return, the group asks for a monopoly over its area of work on the grounds that only those certifiably capable should be authorized to do such work and to define appropriate practice in the area (p. 125).

Both Wise (1994b) and Roth (1994) fear that demand-based policies allowing for uncertified teachers can be devastating to the profession. They argue that such policies are likely to reduce the quality of teaching, lower the livable wage of teachers, and change the resources that are spent on and in schools. In effect, downgrading the importance of certification and training prevents teaching from meeting one of the criteria by which an occupational group calls itself a profession.

[A] shift in locus of preparation [from university programs to alternative certification programs] moves teaching in the direction of trade. On-the-job training is not characteristic of a profession. Dismissing the requirement of professional preparation and a credential prior to practice is also uncharacteristic of a profession (Roth, 1994, p. 267).

This battle over control of training is not new. For over 150 years *who* certifies teachers and *how* that certification is to be done has been a topic of intense debate. At all times, as might be expected, professional educators have fought to control the process, using medicine and law as their models (Angus, 2001).

### **On-the-job Training and Teacher Certification**

To counter the argument that teachers can learn all that is required to be effective on the job, Darling-Hammond (2000), Howey and Zimpher (1994), and others argue that there is an inadequate amount of supervision and training provided to novice teachers by schools. Principals and veteran teachers who could serve as mentors generally do not have the required skills, training, or time to provide novices with quality supervision for on-the-job training. With few exceptions, school districts do not now have access to the additional resources needed for the training of teachers, and it is unlikely that such resources can be obtained. Hawley (1992), articulating the views of many others, claims also that there is a body of subject-matter content and subject-matter method, as well as skills and pedagogical knowledge, that needs to be learned prior to teaching. He and other teacher educators argue that it is unlikely that someone without training in subject matter methods could get in front of a class of students and be a successful teacher. This group of scholars rejects the idea that effective teaching can be learned on the job.

Ordinarily, certification should assure the public that a minimal level of competency has been

achieved, thereby insuring that unqualified people are not practicing the profession. Darling-Hammond (2000) believes the extant data supports that claim. For example, in an analysis of state level data she found the percent of new teachers in a state who were uncertified correlated negatively with performance on six different state assessments conducted by the National Assessment of Educational Progress. These correlations ran from  $-.40$  to  $-.63$ . She found equally large positive correlations for the percent of teachers in a state having regular certification and a major in the field in which they were teaching, again using six NAEP data sets. These correlations ran between  $+.61$  and  $+.80$ . Even for state level data the correlations Darling-Hammond found seem to be quite large, allowing her to assert with confidence that students achieve better when they have certified teachers as instructors.

### **Certification and Competency Testing for Novice Teachers**

There are disputes about what should be covered on teacher certification exams because there is conflict about what is necessary for teachers to know in order for them to be effective. Nevertheless, and despite the problem of defining these areas unambiguously, teacher certification exams currently focus on measuring basic skills, content and pedagogical knowledge (Kearns, 1984).

In 42 states, "Candidates [for certification] are required to pass one or more tests of basic skills, general knowledge, subject matter knowledge, or teaching knowledge" (Mitchell, Robinson, Plake, & Knowles, 2001, p. 44). But critics maintain that the basic skills certification tests "measure verbal and mathematical achievement at about the 10th grade level. And many states set their passing scores so low that virtually anyone can succeed" (Olson, 2000). Even Minnesota, usually one of the highest scoring states in the nation on standardized tests, has a teacher basic skills test that currently passes over 99% of the applicants (Scores needed, 2002). The Arizona basic skills assessment was also of this kind. It had only a 1% failure rate (Crehan, Hess, Lawrence, & Sabers, 2000). This basic skills test was abandoned, in part, because of its low failure rate and also because of adverse impact on some of Arizona's minority group candidates. The low failure rates nationwide suggest that teacher competence testing in basic skills areas is much more about symbolic politics (the need to appear as if standards have been put in place), and a lot less about genuine and systematic attempts to upgrade the quality of the profession.

In the area of pedagogical skills and methods, the test developers often determine what good teaching looks like based on some definition of teacher quality. But teacher quality often is defined as having the necessary knowledge, skills, abilities and behaviors of a good teacher, a circular definition providing little guidance. Moreover, ideas about quality change from one decade to the next, as well as from one test developer to another, and the criteria for measuring teacher quality (the knowledge, skills, and abilities a teacher possesses) is not readily agreed on from person to person (Mitchell, Robinson, Plake, & Knowles, 2001).

The National Research Council (Mitchell, Robinson, Plake, & Knowles, 2001) reviewed past and present definitions of teacher quality and competency, finding that past definitions of teacher quality emphasized teachers' virtue. In its more modern form teachers are still expected to be role models for students, representing high standards of personal behavior and expected to transmit worthy cultural and education values. With no diminishment over time, it is still assumed that effective teachers possess certain personality traits, including enthusiasm, curiosity, and compassion, as well as democratic values. And, as always, effective teachers have been thought talented in teaching the prescribed curricula, an increasing concern after

states have invested heavily in the development of standards and accountability systems.

Today, most definitions of teacher competence from which assessments follow, are the product of three organizations, namely, the National Board for Professional Teacher Standards (NBPTS), the Interstate New Teacher Assessment Support Consortium (INTASC) and the National Council for Accreditation of Teacher Education (NCATE). The National Research Council (Mitchell, Robinson, Plake, & Knowles, 2001) notes that all three sets of standards of quality promoted by these organizations examine teaching in the light of student learning, a relatively new concern. While a focus on learning seems obvious to parents and politicians, this is a much more problematic an issue than it appears to be. The three organizations also acknowledge that teachers' actions or performances depend on many different kinds of knowledge as well as the dispositions to use that knowledge. And they recognize that teachers must also work with others to support the learning and success of all students. The standards of competence described by the three organizations relate to a teacher's commitment to students and their students' learning, requiring that teachers:

- should act on the belief that all students can learn;
- should have deep subject matter knowledge about the substance and structure of their disciplines;
- need to manage and monitor student learning, identify learning goals and choose from teaching styles to meet these goals;
- need to be reflective about their teaching, evaluate their decisions and experiences to make adjustments in their teaching;
- must be part of a larger community consisting of school staff, parents, and the broader non-parent community.

As the National Research Council notes, the standards currently in use "...illustrate the wide range of knowledge, skills, abilities and dispositions that contemporary educators believe competent teachers must possess and demonstrate in the classroom" (Mitchell, Robinson, Plake, & Knowles, 2001, p. 31). Given the wide range and lengthy descriptions of the knowledge, skills, abilities and dispositions that the various assessors of teachers' competence have considered necessary to begin to teach, it seems anomalous that arguments to drop teacher education or to provide emergency certificates would have any credibility at all. Test developers find effective classroom teaching to be extremely complex, requiring the coordination of many different kinds of skills and dispositions, many of which cannot be observed directly. It is hard to imagine that an occupation with these characteristics can be effectively learned on-the-job.

This very same occupational complexity also handicaps the developers of teacher certification testing, leading Sabers to note that, "test developers and researchers haven't done a good job of telling the public that they can't measure everything with a test" (in Crehan, et al., 2000). The public believes that a certification exam will eliminate poor teachers from schools and in essence guarantee that teachers who pass these exams are of high quality. But we do not yet have such tests. At present, it is fair to say that many aspects of teaching cannot be assessed by using a multiple choice or essay exam, and if performance tests of teaching were used such testing would be prohibitively expensive (Crehan, et al, 2000).

## **Validity problems with certification tests**

. All teacher certification exams have problems with construct, content, consequential and criterion-referenced validity (Laczko-Kerr, 2002). In brief, there is no evidence that the construct measured in teacher certification exams is understood. In addition, course content varies more widely across teacher training institutes than, say, law schools or medical schools. This invites criticisms about the content validity for teacher certification exams. Additionally, certification tests limit the pool of potential teachers based on race and ethnic background (Murnane, 1991). As the need for teachers of color increases because of the increase in children of color in our public schools, the numbers of minority teachers seems to be decreasing (Gitomer et al., 1999). One reason for this is the increased requirements for insuring teacher quality, including certification testing for teachers. But these exams have an adverse impact on the teacher supply and this raises concerns about the consequential validity of the exams. Finally, teacher certification exams do not appear to have criterion-related or predictive validity (Smith & Hambleton, 1990; see also Glass, 2002). Certification tests simply do not predict success in teaching. Rather, their intent is to screen out certain applicants from the teaching pool (Sabers in Crehan, et al., 2000). Jaeger, quoted in the National Academy of Science/National Research Council report has an additional concern, namely, that "the sorts of experimental or statistical controls necessary to produce trustworthy criterion-related evidence [are] virtually impossible to obtain" (Mitchell, Robinson, Plake, & Knowles, 2001, p. 72). These problems with validity, particularly predictive validity, seem to bolster the arguments of those who support emergency certification. If the tests cannot predict teaching competency, they argue, why must they be required for certification? The answer offered to that question by supporters of teacher certification tests is that passing the tests ensures familiarity with a broad teacher education curriculum, without which beginning teachers would not do well. This debate can be restated as a version of the simple question we noted above: Does teacher education and the certification that accompanies such programs make a difference in the achievement of students?

## **Research on Certified Teachers and Student Achievement**

Three major areas of research are salient for understanding the importance of certification. First is the research on the effects of certification regarding teachers' content knowledge, particularly mathematics and science knowledge, as it affects student achievement. A second area of research deals with the effects of certification regarding a teachers' pedagogical knowledge, and its effects on student achievement. It is clear that the federal government is having troubles deciding on the relative importance of these two areas, paying lip service to the latter but more often endorsing the former. As "Washington Post" reporter Jay Mathews notes (2002), first the Bush administration pushes through an education bill that demands a "highly qualified" teacher in every classroom. Then the administration releases a report arguing that the nation's education schools spend too much time on classroom methodology. Mathews points out that mixed messages are being sent to the public. But in fact, they aren't very mixed.

Education Secretary Roderick Paige and other Education Department officials claim that schools of education need to spend less time on pedagogical issues and spend more time worrying about whether teachers understand what they teach. The current mantra of federal educational administrators seems to be "You can't teach what you don't know" (Mathews, 2002). Not mentioned explicitly, but implied, is that schools of education should have little role in the training of teachers. Secretary Paige's comments are all the more puzzling from

some one who advocates evidence- based research. His own Department of Education recently requested a review of "rigorous empirical research" on teacher preparation (Wilson, Floden & Ferrini-Mundy, 2002). The authors of this government commissioned report concluded that subject matter knowledge is not sufficient for effective teaching to take place.

[The studies reviewed] suggest that the subject matter preparation....prospective teachers currently receive is inadequate for teaching toward high subject matter standards, by anyone's definition. [Without training in pedagogy] it appears that prospective teachers may have mastered basic skills but lack the deeper conceptual understanding necessary when responding to student questions and extending lessons beyond the basics (p.192).

The third area of research focuses on two sub-areas that both deal with traditional certification and alternatives to it. One of these areas of research is on the effects of regularly certified teachers teaching in or out of their area of expertise. In this literature in-field vs. out-of-field teaching performance is compared, such as when an English teacher is assigned to teach algebra. Out-of-field teaching can be viewed as teaching without the appropriate certification to do so. The second sub-area is concerned with the effects of alternatively certified teachers in comparison to traditionally certified teachers. Present government policy has decided that alternate means to certification are appropriate, with officials claiming that:

[T]here is no evidence that lengthy preparation programs achieve [their] goals any better than streamlined programs that quickly get talented teachers into the classroom....Requiring excessive numbers of pedagogy or education theory courses acts as an unnecessary barrier for those wishing to pursue a teaching career (Mathews, 2002).

Our evaluation of this literature, reviewed in more detail below, is that there is sufficient evidence to conclude that 1) subject-matter knowledge is an important, but not sufficient, factor in a certified teacher's success with mathematics and science students in the upper grades; 2) that teachers who have training in pedagogy outperform teachers without such training; and 3) that traditionally certified teachers teaching in their area of certification outperform both certified teachers teaching out-of-field and alternatively certified teachers. The data on these issues, however, is certainly not unequivocal, and dissenters to all these conclusions exist (see Ballou & Podgursky, 1999; Peck, 1989; Miller, McKenna, & McKenna, 1998). We look at these literatures next.

### **Teacher Subject-matter Knowledge**

. Studies related to teacher subject matter often either evaluate 1) whether a major or minor in a subject area, e.g., mathematics, effects student achievement (Hawk, Coble, & Swanson, 1985; Goldhaber & Brewer, 1996; Monk & King, 1994); 2) whether a passing score on a certification exam provides evidence that certain subject matter has been mastered or that certification affects student achievement (Ashton & Crocker, 1987; Byrne, 1983; Strauss & Sawyer, 1986; Glass, 2002); or 3) whether advanced degrees, e.g., master's degree, or professional development increase student achievement (Goldhaber & Brewer, 1996, 2000; Fetler, 1999; Ehrenberg & Brewer, 1994). Each of these areas of evidence will be reviewed separately.

Researchers have usually, though not always shown that having a major or minor in mathematics or science is beneficial to student achievement in those content areas. Hawk,

Coble, and Swanson (1985) provide research on that issue obliquely, by comparing in-field and out-of-field teaching; concluding that student achievement, for general mathematics as well as algebra, is greater for students who are taught by teachers certified in mathematics (in-field teachers, possessing a major or minor) than is the achievement of students taught by teachers certified in some other content area (out-of-field teachers, neither a major or minor in the area). The researchers hypothesize that the greater success of these in-field teachers' appears to be their greater ability to successfully impart content specific knowledge to students, as compared to their out-of-field counterparts. It is important to note that these studies compare teachers who hold a standard teaching certificate in their subject area (indicating specialized content knowledge training) with teachers who also hold a standard teaching certificate in another subject area (indicating a lack of specialized content knowledge training). The study supports the case for certification in a content area, and suggests deficiencies can be expected among those who are teaching in areas for which they are not prepared.

From their research Goldhaber and Brewer (1996) conclude that "in mathematics and science, teacher subject-specific training has a significant impact on student test scores in those subjects" (p. 206). These same researchers go on to say that their results suggest that it is subject-specific training, not teacher ability that leads to such findings. These authors believe that achievement in technical subjects can be improved by a cessation of out-of-field teaching. The generalizability of these results to the humanities and for teachers in the primary grades is unknown.

Monk and King (1994) also evaluated subject-matter preparation and student performance. In an earlier analysis Monk (1994) had found that there was a "positive relationship between the number of subject-related courses in a teacher's background and subsequent performance gains of these teachers students within the indicated subject area" (as cited in Monk & King, 1994, p. 36). Continued investigation of this phenomena revealed interactions, among them, that, "low-pretest students' performance gains in mathematics were more sensitive to the mean level of their previous teachers preparation than were the high-pretest students" (p. 56). This suggests that lower achieving students will profit more from teachers who are well prepared in their subject matter, than might better achieving students. Thus policies that promote uncertified teachers as the instructors of the poorest and the lowest achieving students, which is the way those policies are usually realized, may be particularly harmful.

Byrne (1983) provides a review of thirty studies that relate teachers' subject matter knowledge, measured by subject knowledge exams or coursework taken, to student achievement. These results were contradictory. A majority of the studies showed a positive relationship (17), while a large number (14) showed that no relationship existed. Byrne does not provide more than a tally analysis of the studies included, which is insufficient given the capabilities of meta-analytic research. A re-analysis using meta-analysis would be helpful.

The National Teacher Exam (NTE) was once used as a measure of subject matter knowledge and was extensively studied. It measured both subject matter content knowledge and pedagogical knowledge. Quirk, Witte, and Weinberg (1973) found only a single study, Lins (1946), in which NTE scores were correlated positively with students' average gain in performance from pretest to posttest. But this finding must be reconsidered in light of the work of Strauss and Sawyer (1986). They analyzed the determinants of average student performance on standardized exams and found that a "1% increase in teacher quality, *ceteris paribus*, as measured by standardized test scores [NTE], is accompanied by a 3-5 % decline in the level of failure or rate of failure of students on high school competency examinations" (p. 41). Simply put, increased scores by teachers on the NTE exams, reflecting increased subject matter

knowledge and increased pedagogical knowledge, decreased student rates of failure. Research on the PRAXIS tests, successor to the NTE, has been conducted by Gitomer, Latham, and Ziomek (1999), and will be reported below.

Teachers' advanced degrees are another indicator of subject matter competency. Goldhaber and Brewer (1996), as part of the study reported above, also found that teachers' degree level is significantly related to student achievement in the area in which the degree was obtained. However, when a general production function model is used, teachers with master's degrees appeared to be no more effective than teachers without advanced degrees. Results varied depending on the statistical models that were used to analyze teacher effects. Goldhaber and Brewer (2000) report that mathematics students who have teachers with bachelors or masters degrees in mathematics have higher test scores than students of teachers without these degrees. They report, however, that there is no comparable impact of degree in science. Fetler (1999) confirms the findings in mathematics, noting that "Schools with more experienced and more highly educated mathematics teachers tend to have higher achieving students" (p. 13). But of course, higher achieving students have access to better schools, and thus these kinds of studies require caution when interpreting them.

Ehrenberg and Brewer (1994), however, report unambiguously that teachers' degree level does matter. "The greater the percentage of teachers with at least a masters degree...the higher black students' scores are [on measures of mathematics, reading, and vocabulary that are associated with the High School and Beyond study]" (p. 10). On the other hand, with a Texas sample, Rivkin, Hanushek, and Kain (2000), note just the opposite, namely, that there "is no evidence that a masters degree raises teacher effectiveness" (p. 3). However, these researchers also report "teacher quality is a very important determinant of the quality of education" (p. 30). Thus the conclusions in this report appear contradictory.

Kain (1998) also studied this issue. He reports that in his earlier research, teachers accounted for at least 7% of total variation in student achievement, indicating that a set of teacher characteristics, including certification and training, affect student achievement.

To analyze the affect of teacher degree on student achievement, Goldhaber and Brewer (2000) used data from the National Educational Longitudinal Study of 1998 (NELS: 88), which provided information about students and their teachers in 10th and 12th grades. Students were surveyed as well as tested on one or more mathematics, science, English/writing and history exams. The researchers found that "having a degree in education has no impact on student science test scores and, in mathematics, having a BA in education actually has a statistically significant (at the 10% level) negative impact on mathematics scores of students" (pp. 138-139). Goldhaber and Brewer's research focused on secondary grades. The complexity of the content taught in secondary mathematics classes is undoubtedly greater than that taught in elementary school, so advanced training in mathematics may be required to effectively transmit that content. That same depth of subject matter knowledge may not be required for elementary teaching.

There is some support for this hypothesis, although research on the effects of advanced degrees and subject matter majors for primary grade students is scanty. One such study, however, was published by the National Center for Educational Statistics (NCES) (as cited in Hawkins, Stancavage, & Dorsey, 1998). These results concerned fourth-grade mathematics students who took the National Assessment of Educational Progress (NAEP). The researchers note that "fourth-grade students who were taught by teachers with an undergraduate or graduate minor in mathematics or mathematics education did not perform better on the 1996 mathematics

assessment than students whose teachers had an undergraduate or graduate major in education" (p. 12). But the NCES report also states that, unlike fourth- grade, eighth-grade students who were taught by a teacher with a major in mathematics outperformed students taught by teachers with majors in education.

One conclusion to be drawn from these data is that a teacher's depth of knowledge in a subject matter influences students' achievement more in the upper grades than the primary grades. But the broader conclusion might be that as the content in a subject matter area becomes more complex, teachers need a much deeper knowledge of that subject matter area to help their students learn at high levels. It is not grade level per se, we think, but the complexity of the ideas to be taught that requires the specialized subject matter knowledge of a teacher. Thus the claim made by TFA, that an individual with a subject matter major from an elite college who elects to teach elementary school without certification is bringing great strength to the schools, may be questioned. It seems probable that after basic college level content is mastered, after some threshold of competency in a subject matter domain is crossed, as it is for most college majors, then pedagogical training for teaching in the elementary grades is more important to success than is content knowledge. Support for this interpretation comes from Rowan, Correnti and Miller (2002), in their study of teacher subject matter competency in the early grades. We look at this issue next.

### **Professional Knowledge/Pedagogical Content Knowledge**

The findings from research that examines a teachers' level of education related coursework and their effectiveness with students is extensive, but often contradictory. We believe, however, that some interpretations of this literature are possible, though we note that better research in this area is possible and badly needed.

Ferguson and Womack (1993) found that the amount of education coursework teachers completed explained about 16% of the variance in teaching performance, as measured by supervisor evaluations; this was more variance accounted for than with teachers' content knowledge, as measured by NTE specialty scores. This research suggests that education coursework is a strong predictor of teaching effectiveness, over and above grade point average in a teachers' major and their NTE specialty scores. In their review of research on this same issue Ashton, Crocker, and Olejnik (1986) also found education coursework to have a significant relationship to teacher performance.

More recent research by Wenglinsky (2002), on the link between teacher quality and student performance, supports the belief that teacher inputs do influence student performance. He notes that the greatest influence on student's achievement comes from classroom practices and the professional development that supports them. Wenglinsky's research indicates that "regardless of the level of preparation students bring into the classroom, decisions that teachers make about classroom practices can either greatly facilitate student learning or serve as an obstacle to it" (p. 7). That is, teacher pedagogical decisions and activities (which are separate from but not unrelated to teacher subject matter knowledge) independently make a difference in student achievement.

Rowan and colleagues (Rowan, Correnti & Miller, in press) reached similar conclusions. These researchers found relatively large effects on young students that could be attributed to teachers, independent of school, social class, previous achievement, and so forth. For any given year, looking at a single score, at a single point in time, teachers accounted for 4% to 18% of the

variance in student's reading and mathematics achievement. This yielded effect sizes of .21 to .42. Across years, looking at student growth, the effects of teachers on students were magnified. Analysis of the teachers' effects on student growth in reading yielded effect sizes of from .77 to .88. The teachers' effects on growth in mathematics achievement were equally impressive, yielding effect sizes of between .72 and .85. The effect of teachers' characteristics on student achievement growth, across time, is roughly three times greater than they are on student achievement status measured at only one point in time. When searching for which teacher characteristics make a difference, these investigators found that the most consistent predictor of young children's achievement was teacher experience. Experience was found to be a much better predictor of student achievement than was subject matter competency. Here again we see the relative importance of pedagogy over content knowledge in influencing the achievement of young children.

What is often not discussed in research reports connecting some teacher quality variable and student achievement is that the great bulk of a teacher's pedagogical training and understanding of beneficial classroom practices is provided in their teacher training programs. Clearly experience matters; but that means that preparation to profit from experience must matter as well. And that suggests that the experience gained from intensive student teaching, over a sufficient time period, might also matter. Such experiences are provided as a matter of course in most traditional teacher certification programs, and are missing from most alternative and emergency certification programs. Without adequate teacher training, then, emergency certified teachers and other under-certified teachers could retard student learning as they engage in teacher learning. We examine teacher experience in more detail next.

## **Teacher Experience**

Teacher experience is another teacher quality variable that influences student learning and is indirectly related to issues of certification. Hawkins, Stancavage, and Dorsey (1998) report that in the 1996 AEP analysis, students who were taught by teachers with less than 5 years of teaching experience performed below the level of those students whose teachers had 6-10 years or 25 or more years of experience (p. 22). Fetler (1999) also supports the finding that number of years teaching is positively related to student test scores. Lopez (1995), using a large data set from Texas, reports that teachers require about 7 years of experience in order to be able to maximize their students' test performance. Similarly, Rivkin, Hanushek, and Kain (2000) report that there are small but significant relationships between student achievement gains and teacher experience: "The teacher experience investigation concentrates on entering teachers and supports the notion that those in the first two years of experience do worse than more experienced teachers. New teachers' average student gains are lower by roughly 20 percent of a standard deviation in both 4th and 5th grades" (p. 27). They also report that 4th and 5th grade student achievement on the Texas Assessment of Academic Skills (TAAS) is effected by overall teacher experience. The results indicate a small but significant relationship between achievement gains and teacher experience. The study of Rowan et al., (in press) cited above, supports this conclusion as well.

The point of this discussion about experience is that virtually all university teacher certification programs include both early field experiences and student teaching in their curriculum to provide experience to novice teachers. While we have little empirical evidence to determine what kinds and amounts of experience are the most beneficial, it seems likely that teacher induction programs that offer little or no experience will be deficient. This is a criticism of the TFA program and any other programs supporting emergency or alternative certification that

allow un-experienced and inexperienced teachers to become classroom instructors. Reviewing similar literature for the Department of Education, Wilson, Floden and Ferrini-Mundy (2002) conclude that the parts of the teacher education experience that make a difference are unknown, but that "the research suggests that there is value added by teacher preparation (p. 194)." They also state that clinical experiences and field-work, such as that provided through student teaching, are seen as powerful forces—maybe the most powerful force—in programs of teacher preparation.

Interestingly, if a state policy provides for emergency certification to teach for only a short period of time, they may do a disservice to students, since it is through experience that teachers acquire their competency. The logic is this: It may be wrong to employ emergency certified teachers, but to dismiss them solely on the basis that they served two years, the maximum for an emergency certificate in some states, is to negate and reject how much they may have learned in that time. On a case-by-case basis, it may be better to decide if an emergency teacher has been reflective about his or her experience and thereby learned enough to be effective. It may compound the original error to dismiss them after a short period of time.

The review of research on content knowledge, pedagogical knowledge, and experience, given above, focuses on where these abilities and characteristics of teachers are to be learned, and in what mix, but there is no major dissent about their importance for student learning. Wenglinsky (2002) makes this case best using data from the National Assessment of Educational Progress (NAEP) to examine the role teachers and their instructional practices play in influencing student achievement. He summarizes his findings as follows: "The study finds that the effects of classroom practices, when added to those of other teacher characteristics, are comparable in size to those of student background, suggesting that teachers can contribute as much to student learning as the students themselves."

### **Appropriate Assignment of Certified Teachers**

Some research on certification status supports the fact that teachers who are certified and teaching in the area in which they are certified outperform teachers who are less than fully certified and teachers who have no certification (Darling-Hammond, Berry, & Thoreson, 2001; Darling-Hammond, 1992; Fenstermacher, 1992; Evertson, 1984). Unfortunately, however, not all certified teachers are assigned to teach in the areas for which they have been trained (Goldhaber & Brewer, 1996; 2000; Ehrenberg & Brewer, 1994). In fact, large numbers of the certified teaching staff are assigned duties for which their certification is irrelevant. This kind of out-of-field teaching is exemplified in the elementary grades when a fifth grade teacher is assigned to teach a second grade class, or in high school, when an English teacher is assigned to teach an American History or a biology class. The most cited reason for out-of-field teaching is a lack of fit between the teachers on the staff and the teaching assignments that an administrator must make (Ingersoll, 1999a).

In some curricula areas such as bilingual and special education, science, and mathematics, there is a teacher shortage. This appears to be caused by increased student enrollments, retirements due to the "graying" of the teaching profession, increased turnover by teachers due to increased difficulties in carrying out their jobs, and the increased remuneration for mathematicians and scientists for work in other fields. If faced with difficulty filling positions, school boards and administrators think of three solutions: they hire less qualified teachers, they assign teachers trained in another subject area to teach in the understaffed areas, or they make extensive use of substitute teachers.

There are many problems with the teacher shortage explanation for out-of-field teaching. First of all, this explanation does not provide an answer for why large numbers of out-of-field teachers exist in fields like English and social studies, two areas of teaching that have long had a surplus of teachers. Another problem with this explanation is that it has only been within the past few years that schools have had trouble filling teaching openings, while the problem of out-of-field teaching has been with us for a significantly longer period of time. Finally, the teacher shortage is based on the assumption that there are too few teacher candidates. But in fact, the biggest difficulty is that teachers are choosing to leave their profession for other jobs (Ingersoll, 1999a). Ingersoll (1999a) comments, "Out-of-field teaching is common, I believe, because it is not only legal but also more convenient, less expensive and less time consuming than the alternatives" (p. 7).

Ingersoll (1996) evaluated data from the School and Staffing Survey to determine the proportion of teachers who teach out-of-field. He provides data indicating that one-fifth of public school students enrolled in 7th-12th grade English courses are taught by teachers who did not have at least a minor in English or some other closely related field. Of the students enrolled in 7th-12th grade mathematics courses, more than a quarter are taught by teachers without a minor in mathematics or mathematics education. The results are less drastic in other areas of course work. In Arizona, 35% of teachers in English, 39% of teachers in Math, 35% of teachers in Social Studies and 27% of teachers in Science were assigned to teach secondary courses without a major in that subject area ("Quality Counts," 2000).

Research also supports the belief that out-of-field teaching is related to levels of school poverty (Ingersoll, 1996; 1999b; Haycock, 2001). Ingersoll (1996) reports, "in no fields did high-poverty schools have less out-of-field teaching than did low poverty schools, while in several fields, students in high poverty schools received distinctively more out-of-field teaching than in low poverty schools" (p. 5). This trend is similar for students who are placed within different educational tracks in their courses. High track students are exposed to less out-of-field teaching than low track students (Goldhaber & Brewer, 2000; Rivkin, et al, 2000, Ingersoll, 1996), while "minority and poor students are disproportionately placed in lower track and lower achievement courses, [which] critics claim are taught by the least qualified" (Ingersoll, 1996, p. 1). Darling-Hammond (1997b) reports that in the most heavily minority schools and inner cities less than 50% of the teachers in mathematics and science are licensed and have a degree in the subject they teach. Darling-Hammond remonstrates that throughout the country we have the least qualified teachers teaching the most disadvantaged students, while the most qualified teachers are teaching the most advantaged students.

At the secondary level the relationship between in-field teaching and student achievement is stated forcefully by Hawk, Coble, and Swanson (1985). They conclude that:

In field certified math teachers know more math and show evidence of using more effective teaching practices than their out-of-field counterparts. Further, and most important, students of in field certified math teachers achieve at a higher level than do students taught by out-of-field teachers (p. 15).

In short, a certified teacher teaching in the field for which they were prepared performs better than when assigned to areas for which they were not prepared. Preparation matters.

### **Alternative Routes to Teaching**

Much of the research that draws attention to alternative certification programs does not adequately address the issue that many such programs are similar in both the level and rigor of training provided by traditional certification programs (Buck, Polloway, & Robb, 1995; Miller, McKenna, & McKenna, 1998; McKibbin, 1988; Bliss, 1992; Stoddard, 1992; Darling-Hammond, Berry, & Thoreson, 2001). On the other hand, many of the alternative teacher training programs are poorly designed and administered, providing little in the way of appropriate training (Wilson, Floden, Ferrini-Mundy, 2002). Although increasing dramatically in number, there are currently no standards for assessing alternative certification programs. The large variability in alternative certification programs makes research on this phenomenon difficult. (Of course, to be equally fair, we must note the variability in traditional programs of teacher education, whose design and administration have also been noted by many to be equally slipshod. Even accredited programs have, in our opinion, some embarrassing design characteristics.)

Advocates of alternative certification, however, claim that they provide teachers for urban and rural schools and in specific shortage areas, i.e., mathematics and science. Zumwalt (1991) summarizes research on several alternative certification programs and reports that they do attract teachers who are more willing to work in rural or urban poor districts than traditionally trained teachers. McKibbin and Ray (1994) also report that alternative certification programs attract people with subject matter majors like mathematics and science who are interested in teaching, but not interested in traditional teacher certification.

Alternative certification is also seen as a cost-effective way to train people who did not or will not enroll in conventional undergraduate or graduate education programs. Such programs are cheaper (Zumwalt, 1991), as might be expected from programs that are shorter in duration and provide less instruction, supervision and assessment of their students.

Proponents of alternative certification claim that these programs attract better quality candidates who are more academically able than those who attend traditional certification programs (Kanstroom & Finn, 1999). Participants of these programs are generally people who have majored in traditional academic subjects rather than education. It is a major assumption of alternative programs that subject matter content knowledge is more important to teaching than is education related coursework (Jelmberg, 1996). But some studies show that the teachers in alternative routes to certification have high drop-out rates from both the programs of instruction and from actual teaching. (Wilson, Floden & Ferrini-Mundy, 2002). Other studies show that alternative certification recruits in mathematics and science have lower grade point averages than recruits in traditional teacher education programs (Stoddart, 1992). Moreover, to date, these alternatively certified teachers have not demonstrated strong skills in their content area. Furthermore, teachers from alternative routes to education, including TFA teachers, when compared to those trained in more traditional teacher education programs, report many more problems with their preparation programs. For example, on 39 of 40 different questions the TFA teachers rated their preparation more poorly than did those who were trained in more traditional programs. The self-confidence and sense of efficacy of those prepared in traditional programs was higher than for those who came to teaching through alternative programs (Darling-Hammond, Chung, & Frelow, 2002).

When these facts about alternative routes to teaching are added to research that debunks the belief that subject matter knowledge is more important than education related coursework (Darling-Hammond, 2000; Gitomer, Latham, and Ziomek, 1999; Monk and King, 1994; Wilson, Floden & Ferrini-Mundy, 2002), we see evidence of an educational policy that must be seriously questioned.

Proponents of alternative certification also make the claim that traditional certification programs attract mostly twenty-something, white women, a problem because the students of the public schools in urban areas are not often white, and some commonality and similarity in life experience is probably a better background for teaching positions in these more difficult settings. Proponents of alternative certification correctly note that they often attract a more diverse group of candidates, specifically men, older adults, minorities and retired military personnel (Bradshaw & Hawk, 1996; Eifler & Potthoff, 1998; Hawley, 1992; Houston, Marshall, & McDavid, 1993; Keltner, 1994; Kennedy, 1991; Kwiatkowski, 1999); Natriello & Zumwalt, 1993; MacDonald, Manning, & Gable, 1994; and Stoddart, 1993). Wendy Kopp (1994), Teach for America founder, notes that short alternative certification programs allow young adults who are unsure of their career options an opportunity to serve students.

Alternative certification may actually be an improvement over simple emergency certification, which allows almost anyone with a bachelors degree to teach without any preparation to speak of. But some researchers (Bradshaw & Hawk, 1996; Berry, 2001) are critical of the level of professional knowledge demonstrated by alternatively certified teachers. Alternatively certified teachers tend to have a limited view of curriculum and a lack of understanding of student ability as well as motivation; they experience difficulty translating content knowledge into meaningful information for students to understand; they are less effective planners of instruction; and they tend not to learn about teaching through their experiences. Research is also critical of the supervision and mentoring support that is given to preservice teachers in most alternative certification programs (Smith, 1991; Darling-Hammond, 1992; McKibbin & Ray, 1994; Bradshaw & Hawk, 1996; Jelmberg, 1996; Miles-Nixon & Holloway, 1997; Berry, 2001). It appears likely, from the extant research and criticism, that poorly run and short alternative certification programs, particularly those that do not provide much classroom experience and supervision, may actually not be any better than simply hiring emergency certified teachers with no teacher education experiences.

### **Teach for America**

The most familiar of the alternative certification programs is Teach for America (TFA). This ambitious program recruits graduates from top universities, provides them minimal training, and places them in public school classrooms across the nation to teach. The public schools, however, are all in either rural or poor urban districts (Darling-Hammond, 1994). Research conducted on TFA has been less than encouraging.

Four separate evaluations found that TFA's training program did not prepare candidates to succeed with students, despite the noticeable intelligence and enthusiasm of many of the recruits. Most criticism of a corps member's teaching behavior (classroom management was the greatest area of concern, followed by insufficient knowledge of the fundamentals of teaching and learning) was qualified by the cooperating teachers' perceptions of limitations of the program in providing the corps member with adequate practice or theory to be successful (Darling-Hammond, 1997a, p. 310).

From an interview study by Stevens and Dial (1993), TFA teachers apparently decide to teach because they like working with children; they didn't have other options; and they felt that TFA was their best alternative given their "circumstances and indecisiveness at the time" (p. 70).

Jonathan Schorr (1993), a former TFA teacher, describes the inadequate training and preparation that he and other TFA teachers received prior to being placed into schools. He notes, "just eight weeks of training ... is not enough for teachers" (p. 316). Schorr admits, "I was not a successful teacher, and the loss to the students was real and large" (p. 318). Schorr offers the first-hand experience that makes Darling-Hammond (1994; 1997a; 2001) quite critical of TFA, specifically due to the program's limited training of candidates, lack of evaluation, and the fact that such a program perpetuates the placement of poorly trained teachers with the most needy students in the nation.

Raymond, Fletcher, and Luque (2001) conducting research for the Center for Research in Education Outcomes (CREDO), released a report evaluating the Teach for American program in Texas. The report compares scores on the Texas Assessment of Academic Skills (TAAS) of students taught by TFA teachers and non-TFA teachers, and lauded the performance of TFA teachers. However, the National Commission on Teaching and America's Future released a response outlining serious concerns with the research (see Darling-Hammond, 2001 and 2002). The most important of the criticisms is that the performance of the TFA teachers was never compared to the performance of regularly certified teachers. The comparison used to assess the TFA teachers was other uncertified teachers, some of whom didn't even have four-year college degrees.

It should also be noted that when we tried to access the data for this report, we were informed from both the primary researcher and the Texas school district responsible for the data that it was not available for independent review. We were told that the data was not the property of the researchers who reported the study, nor did it belong to the district, and that neither had a complete data set to provide for independent analysis. In our opinion, therefore, it is appropriate to regard this report as irrelevant, given that the comparison used to assess TFA teachers was faulty, the data are not available for verification or replication by other scientists, and the report has not been published in a peer-reviewed journal.

### **Traditional Teacher Certification**

In this section we discuss traditional programs of teacher certification, with the understanding that they vary enormously, as do the alternative certification programs. The Holmes Group (1986), active in teacher education reforms, reports that competent teaching consists of three elements: subject matter knowledge, systematic knowledge of teaching, and reflection on experience. Members of the Holmes group defended teacher education in the universities by arguing that they do, in fact, prepare people to successfully integrate these three elements into their professional lives. In their defense we note that teachers entering the field from university teacher education programs are generally more academically able than the average college student, while unlicensed entrants into teaching have significantly lower levels of academic achievement than most college students and were lower also than those who were prepared by the university to teach (Gitomer, Latham, & Ziomek, 1999). In fact, "reviews of research over the past thirty years, summarizing hundreds of studies, have concluded that even with the shortcomings of current teacher education and licensing, fully prepared and certified teachers are better rated and more successful with students than teachers without this preparation" (Darling-Hammond, 1997a, p. 308; Evertson 1984). Apparently disregarding this research, Former Secretary of Education Chester Finn proposes that the common sense route to boosting teacher quality is to simplify entry and hiring. Fundamentally, he argues, we should let anyone who wants to teach do so, and simply deregulate the teacher certification process (Kanström &

Finn, 1999; Finn & Kanstroom, 2000). Finn is also the president of the foundation that helped support the unverifiable Raymond, Fletcher, and Luque (2001) study that is so supportive of uncertified teachers. Nationally syndicated conservative columnist Thomas Sowell supports Finn's position. Sowell (2002) says bluntly that college of education courses are "drivel" and falsely reports that the academic record of those who enter teaching through the university route is deficient in comparison to almost all other college majors.

The research martialed in support of prepared and certified teachers includes research demonstrating that teachers hired without preparation or only partial training experience difficulty meeting the needs of the students in their classrooms. Such individuals have more difficulties than fully prepared teachers do in accomplishing their day-to-day job requirements (Darling-Hammond, 1997a). Prospective teachers apparently perform better if they have completed a teacher preparation program focused on content knowledge, pedagogical course work (including learning theories, developmental theories, theories of motivation and issues of student assessment) and practice teaching. Although variations in the philosophy, implementation and quality of teacher education programs are enormous, research nevertheless suggests that many versions of this form of preparation are successful in providing adequately trained teachers for the complexity of classroom instruction (Ashton & Crocker, 1987; Darling-Hammond, 1992; Wilson, Floden, & Ferrini-Mundy, 2002).

McDiarmid and Wilson's (1991) research is relevant to this point. They demonstrated that subject matter knowledge is not sufficient to prepare teachers for teaching the concepts in these fields to students. They did this by evaluating mathematics majors in alternative certification programs that stressed subject matter knowledge and found that those teachers had strongly held misconceptions about the subject matter and the appropriate ways to teach the content. Their results indicated that these teachers were unable to integrate their subject matter knowledge with teaching practices to allow for effective instruction. In effect, because they were lacking in education coursework, they were unable to provide the appropriate instruction to students.

Another aspect of good quality certification programs is that they provide experiences for the preservice teachers in classrooms both under direct supervision and with continued mentoring. Darling-Hammond (1992) notes that skills need to be learned in context, where they can be practiced under supervision. The student teaching experience allows the preservice teacher to transform information from coursework in order to learn its character in the context of the real world of teaching in classrooms. Jelmberg (1996), cited above, compared traditionally certified teachers with alternatively certified teachers' performance based on their experience. His results showed that experienced teachers from traditional certification programs are rated higher on instructional skills and planning by their principals, and perform better, than did experienced teachers who came from alternative certification programs.

### **Emergency certification**

The reviews of research, above, compared alternative certification programs, some of which provide enough preparation for graduates to receive full certification, while others provide minimal training resulting in graduates receiving either a provisional or emergency certificate. Research reviewed above also compared fully certified teachers to one another, distinguishing between teachers who taught in-field or out-of-field. There is little research comparing emergency certified teachers and regularly certified teachers.

Emergency certificates are issued to prospective teachers who have met some, but not all of the requirements for state certification. Minimum requirements are often a bachelor's degree and a passing score on a certification exam. Emergency certificates are issued for a limited time period, usually one to two years. Some states allow for these to be renewed, while others states issue a one time only, nonrenewable certificate. In 1998 data about certification waiver rates were available from 39 states (U.S. Department of Education, 1999). Sixteen of the 39 states had waiver rates greater than 2% of their teaching population, eight had rates over 5% while some had rates as high as 17%.

Emergency certified teachers are more likely to be hired in already low performing schools, schools that serve low SES students, schools in rural and inner city areas, and for positions that are hard to fill (Darling-Hammond, 1997a; Darling-Hammond, 2001; Ingersoll, 2001; Gitomer, Latham, & Ziomek, 1999). Since high poverty schools are more likely to have high rates of out-of-field teaching as well as more emergency certified teachers, it may be that some states are failing to provide the "adequate education" that most state constitutions require (Ingersoll, 2001; Hirsh, Koppich & Knapp, 2001; Rivkin, et. al., 2000). And now that federal government has waded in on this issue, requiring a highly qualified teacher in every classroom, the competency of teachers with emergency certification is sure to be questioned further.

Emergency certification is justified on the basis of three arguments. First, it is argued, that there is a teacher shortage requiring that states emergency certify teachers to provide enough teachers for every classroom. But the shortage may not be as severe as many claim (Feistritzer, 1994; Ingersoll, 1997, 2001; Ballou, 1996; Hirsh, Koppich, & Knapp, 2001; Hardy, 1998). The National Center for Education Statistics indicates that the teaching force will grow to more than 3 million teachers by the year 2008 (US Department of Education, 1999). But it is a misconception that colleges of education will need to train millions of new teachers to meet the needs of school districts. Darling-Hammond (cited in Hardy, 1998) believes that this potential shortage is not universal, claiming that "there are districts that experience difficulty hiring qualified teachers, but overall, we have a surplus of teachers" (p.20). The teacher shortages are seen in subject areas like mathematics and science; in the service areas for special needs populations, such as special education and bilingual education; and shortages exist in rural areas and in inner city school districts (Wayne, 2000; Natriello & Zumwalt, 1993; Hardy, 1998; Hirsch, Koppich, & Knapp, 2001).

The projected teacher shortage is also based on assumptions of increased student enrollments and an aging workforce. But these assumptions have similarly been questioned. Research evaluating the Survey of Recent College Graduates (Ballou, 1996) has shown that:

In every year there were at least twice as many [qualified] applicants as there were persons hired in full-time public school positions. Far from indicating that the nation faces a teacher shortage, these data show that the teacher labor market as a whole has been in a chronic state of excess supply, though shortages may arise in specific locations and subject areas (p. 101).

Research also indicates that regularly certified teachers are in short supply because of poor pay; low levels of job satisfaction, particularly when working with disadvantaged minority students (Hanushek, Kain & Rivkin, 2001); and because limited faculty input about the management of schools discourages college graduates from teaching and drives current teachers out of the profession. Ingersoll (2001) suggests that:

The imbalance of teacher supply and demand at the root of school staffing problems is neither synonymous with, or primarily due to, teacher shortages in the technical sense of a deficit in the quantity of qualified candidates. Rather than insufficient supply, the data indicate that school staffing problems are primarily due to excess demand, resulting from a 'revolving door'—where large numbers of teachers depart their jobs for reasons other than retirement. Thus the solution...does not primarily lie in increasing supply, but rather in decreasing demand (p. 501).

A second argument for emergency certification is that there are many people who would teach, but do not, because standard certification requirements prohibit them from doing so. Thus opponents of traditional teacher education programs call for the removal of certification requirements, claiming that there is no "special body of knowledge" that teachers need to know in order to be successful. These advocates for the abolition of requirements claim that what needs to be learned by new teachers can be learned in the first year of teaching (Roth, 1994; Kanstroom & Finn, 1999; Finn & Kanstroom, 2000). In fact, anecdotal evidence claiming that teachers learn more in the first year of teaching than they do in their education courses is quite strong. Armed with this knowledge, it is then argued that a person who holds a college degree, in possession of some level of content knowledge, and with some limited experience teaching youth, is competent enough to begin to teach. Such beliefs drive the movement against certification despite research that argues against this position (McDiarmid & Wilson, 1991).

Traditional certification programs are rejected by many adults who would be interested in teaching as a second career but who will not or cannot engage in time consuming and expensive regular teacher certification programs. Proponents of alternative certification believe that these adults, called career transitioners, have skills that they have learned in their other employment that could be used to teach children (MacDonald, Manning, & Gable, 1994). Additionally, some believe that adults have unique life skills and experiences that can be useful to students (Zumwalt, 1991). Research does indicate that alternative certification programs attract an older and more diverse population, through their more flexible schedules, less stringent requirements, and so forth; however, it is unclear from this research that certification needs to be waived in order to recruit a more diverse teaching population with many life skills and employment experiences (Natriello & Zumwalt, 1993; Bradshaw & Hawk, 1996; Hawley, 1992).

The third argument for emergency certification makes use of the long-standing lack of confidence by state officials and the general public in the quality of the teachers who graduate from colleges of education. Too often colleges of education are perceived to attract less able students, thus producing under qualified teachers. This is simply not true. Research supports the assertion that the academic quality of students entering colleges of education is quite good. For example, Gitomer, Latham, and Ziomek (1999) showed in their analysis of ACT and SAT scores that students from colleges of education were as academically skilled as students with other college majors. They also reported that traditionally certified teachers have the highest passing rates on certification exams (PRAXIS I and II) compared to alternatively certified and emergency certified teachers, even though they appeared to be similar in initial achievement, based on SAT scores. They concluded that traditional certification (having training in teaching methods, pedagogy and practice in teaching) makes a difference on licensure. They attribute the better performance of traditionally certified teachers to the training and instruction, provided by colleges of education.

## Summary of Related Research

With regard to teacher subject matter the current research suggests to us that in mathematics, especially, and at the upper grades, in particular, subject matter competency as assessed by college majors, courses taken, and degrees held, leads to higher student achievement. Professional pedagogical knowledge appears equally important a contributor to student achievement at the upper grades, and may even be more important than content knowledge in the elementary grades.

With regard to experience through teacher education course work and by means of learning on the job, the research suggests that student achievement is affected in positive ways. The powerful effects of content knowledge, pedagogical knowledge, and experience, contribute to success in teaching in one's own field. When teaching out-of-field, such knowledge and skill is of less use and teacher effectiveness is compromised.

When we examined alternative routes to teaching we found them to be quite variable. Still, in comparison to traditional teacher education programs (that are also quite variable) the alternative programs take less time, are less costly, attract more diverse students, but also record higher drop-out rates. Because they usually take much less time, alternative programs may suffer from under-preparing students for the classroom, a problem compounded by the lack of supervision and support given by the hiring districts. Teach for America, as one alternative route to teaching, appears to suffer from the ordinary and typical problems inherent in the designs of such programs. While criticisms abound, there is a substantial body of literature suggesting that traditional teacher education programs, warts and all, seems to provide more competent appearing teachers whose students perform better.

The research suggests that emergency certified teachers are probably the least prepared to do well. Unfortunately, such teachers are typically hired into districts with the hardest to teach students.

After reviewing this literature it seemed as if the conditions necessitating out-of-field teaching and the hiring of alternate and emergency certified teachers could easily be modified, eliminating the need for these practices to exist. In the past, however, neither politicians nor school boards had the will to do that. But under the new federal *No Child left behind Act of 2001*, school districts will have to have a "highly qualified" teacher in every classroom or lose federal funding. It will be an interesting few years as ways to define "highly qualified" and the related term "competent" are worked on to meet the letter of the law. The research in these areas is also likely to be reinterpreted in light of that law. In the near future, surely both the definitions of, and the research associated with the idea of teacher quality, will find their ways into courtrooms of America. This study may help the courts in thinking about what it means to have a highly qualified teacher in the classroom.

## Method

### Research Design

An ex-post-facto archival research design was used to study the performance of students in the classes of the under-certified and certified teachers in selected districts of the state of Arizona. Districts provided both the information about the teachers participating in this study and their

Stanford Nine (SAT 9) class means. The SAT 9 data provided by districts were also compared with the same data obtained from the Arizona Department of Education. SAT 9 data were available from 1997-2000, but was not available for the 2000-2001 academic year.

## Sample

In Arizona under-certified teachers have three labels: "emergency" (for holders of bachelor degrees from accredited institutions, with little or no educational coursework, who can get clearance of criminal background through fingerprint analysis); "temporary" (a rarely used designation similar to "emergency"); and "provisional" (for those with some, or even considerable teacher education training, who are short certain units or requirements that could earn them a standard certificate). In opposition to the under-certified teachers are the regularly certified teachers who meet all of the state requirements for certification. These teachers hold a bachelor's degree from an accredited institution, have completed 45 semester hours of elementary or secondary education course work, obtained a passing scores on the AEPA, demonstrated that they understand both the Arizona and US constitution, and been vetted for any criminal background through analysis of their fingerprints. (The definitions of under-certified and certified teachers, given above, were accurate through the year 2000, when the Arizona Department of Education made changes to the certification laws. Certification requirements are still in flux, and so current Arizona certification policies may not be the same as those reported in this study.)

Among the under-certified teachers in this study are some from an alternative route to traditional teacher training, the program "Teach for America (TFA)." Teach for America is a popular alternative certification program. Its mission calls for placing energetic, bright, but unqualified teachers into poor, urban school districts (Darling-Hammond, 1994). The instructional effectiveness of under-certified teachers in general, and the TFA teachers in particular, is of considerable interest to the policy community.

*District Selection.* Arizona school district superintendents listed by the department of education as participating in the Federal Teacher Shortage Loan Deferment Program (Arizona Department of Education, <http://ade.az.gov/researchpolicy/ts>) were invited to participate in this study. This federal program requires the State Department of Education to generate a ranking of school districts in the state by the percent of under-certified teachers in each district. This list of school districts provided a convenient population from which to obtain a sample. Of the school districts receiving a request to participate (N=24 for the 1998-1999 dataset, N=12 for the 1999-2000 dataset) five school districts responded positively. These five school districts represent 20.8% and 16.6%, respectively, of the school districts designated by the state as having severe teacher shortages. All five school districts chosen for this study were included in the Department of Education's 1998-1999 classification of school districts. Only two of these school districts were also included in the Department of Education's 1999-2000 classification of school districts with severe shortages. All of the participating school districts shared similar characteristics. They all serve inner city student populations, largely minority, and all participate in the Teach for America (TFA) program. All the schools in these districts have difficulty filling teaching positions.

*Methods of data collection.* The five participating school districts provided lists of new hires for the 1998-1999 and 1999-2000 school years. In some cases this list contained information about the teachers' certification status, while in other cases further research was necessary to obtain this information. We were granted permission to access individual teacher personnel

files in order to collect the necessary data on the school where the teacher was employed, the grade level taught, the teachers' certification status, their highest degree earned, the date and institution where their degree was earned, their age (determined from year of their birth), and teaching experience.

Of the teachers included in the dataset, 64% had no prior teaching experience, a judgment based on their hire date, employment history, resume and application. The majority of newly hired teachers were recent graduates from college. Some of the new teachers, however, had delayed their entrance into teaching for many years after graduation, but they had no prior teaching experience indicated in their personnel files. Teachers were removed from the sample if they taught kindergarten, first grade, art, music or special education, grades and subjects not assessed by the Stanford Nine (SAT 9).

The assessment departments of each school district provided test scores aggregated at the classroom level. Included were the teachers' class totals as raw scores, scaled scores, grade equivalent, national percentile rank, stanine, national normal curve equivalent, as well as class percentile rank and class stanine. Individual student scores were not needed or provided. Additionally, state level SAT 9 data was later obtained from the Research and Policy division of the State Department of Education in order to confirm the accuracy of the SAT 9 data collected from the school districts. State level data was aggregated, by teacher name, for each of the school districts. In the event of a discrepancy between the two sets of data, we opted to use the state level data. This discrepancy occurred with only one of the school districts data files.

### **Matching Procedure**

In the five districts studied, 293 teachers were hired in 1998-1999 and 1999-2000 who met the requirements for inclusion in this study. In order to participate teachers personnel files were required to contain the demographic data necessary for analysis, as well as classroom level SAT 9 scores. Teachers' data files were matched using SPSS procedures to sort the files. Teachers in each district were first categorized on the basis of their certification status, under-certified teachers (labeled emergency, provisional or temporary certified teachers by their districts) constituted one group, certified teachers made up the other group. Teachers from each group were then matched based on grade level and highest degree attained. Teachers for whom no matches could be found were removed from the analysis.

Matches were made using the following rules: 1) matches were first made within the same school, 2) matches were made within the same school district, and 3) matches were made between similar school districts. The first and second matching rules serve to minimize exogenous variables associated with student achievement scores, e.g., socio-economic status, school characteristics, curriculum, etc. It is assumed that teachers in the same school teach similar students, an imperfect but reasonable assumption. The identical assumption can be made about schools within the same district boundaries, since Arizona school district boundaries are based on relatively homogenous geographic areas. Conversations with district personnel, in the course of collecting the data at the district offices, provided no evidence that the certified or under-certified teachers were "tracked" in any way. The assignment of teachers to schools, and to classrooms within schools, appears to have been unbiased. Similarly, we have no reason to believe that class size or student ability was different in any way for the certified or under-certified teachers in our sample.

The cross-district matching of teachers, however, is more problematic than the within district matching. We made these matches based on the "sameness" of the two districts. Sameness was determined using data collected from the Education Finance Statistics Center, a subdivision of the National Center for Education Statistics (U.S. Department of Education, 2001). Using public school district financial records for the latest year available, 1996-1997, data about each of the participating school districts were collected for: student teacher ratio, administrative ratio, median income, percent of householders with high school graduates, percent of non-white children, percent of limited English proficiency, and percent of children in poverty. School districts that shared similar characteristics were matched to one another. This procedure matched the Osborn Elementary School District to the Creighton Elementary School District, and matched the Roosevelt Elementary School District to both the Nogales Elementary School District and the Murphy Elementary School District. These are not pseudonyms, but the actual names of the Arizona school districts that generously helped us in this study.

Synthesizing data from the various state departments of education, GreatSchools.net also provided data that were used to match schools having similar characteristics. Sameness matches were also made based on: AIMS reading and math scores, SAT 9 reading and math scores, teacher experience, percent of teachers with masters degrees, attendance rate, open enrollment, percent free and reduced lunch, and ethnic breakdown within the school district. These data support the matches that were made using the NCES finance data with the exception that the Murphy School District was found to more closely match the Creighton School District. Teacher matching across districts, therefore, was accomplished by finding similar school districts serving similar student populations with similar economic bases. The assertion that teachers in these different school districts are sensibly matched is well substantiated, but we acknowledge that the nature of the matching procedures used constitutes a potential flaw in studies such as this. Random assignment of under-certified and certified teachers to classes within districts would have provided a stronger design, but this was not possible. It is important to note, however, that the matches in this study were made with out any knowledge of the teachers' class scores on the SAT 9.

Descriptive analyses were conducted on the complete sample of teachers in the data set (N= 293). These results are provided in Table 1.

**Table 1**  
**Descriptive Statistics for 1998-1999 and 1999-2000 Datasets**

	<b>Certified Teachers N=159</b>	<b>Emergency Certified Teachers N=89</b>	<b>Temporary Certified Teachers N=19</b>	<b>Provisionally Certified Teachers N=26</b>
<b>District</b>				
Murphy	21	13	10	1
Creighton	64	34	0	0
Osborn	21	2	2	2
Roosevelt	41	35	19	19
Nogales	12	5	4	4
<b>Grade</b>				

2	25	9	3	7
3	19	8	2	5
4	29	7	2	4
5	24	13	1	2
6	21	22	4	5
7	22	14	3	2
8	19	16	4	1
<b>College</b>				
ASU	74	13	6	9
U of A	9	6	1	2
NAU	16	5	3	3
U of Phoenix	0	4	0	0
Ottawa	8	0	2	1
Grand Canyon	0	1	0	1
Out-of-State (large public)	14	25	2	4
Out-of-State (small public)	13	8	1	1
Out-of-State (large private)	4	7	0	0
Out-of-State (small private)	6	8	2	4
Foreign Education	9	10	0	0
Prescott College	1	2	2	0
Other	4	0	0	0
	<b>Certified Teachers N=159</b>	<b>Emergency Certified Teachers N=89</b>	<b>Temporary Certified Teachers N=19</b>	<b>Provisionally Certified Teachers N=26</b>
<b>Major</b>				
Education	9	7	1	0
Elem. Ed.	74	7	14	19
Second. Ed.	5	1	0	0
Early Child.	4	0	0	0
C & I	6	0	0	0
Ed. Admin.	6	0	0	0
Bilingual Ed.	8	0	2	0
Spec. Ed.	2	0	0	0
Phys. Ed.	1	0	1	0
Liberal Arts	10	12	0	0

Soc. Sciences	23	25	11	4
Sciences	4	17	0	3
Business	4	7	0	0
Languages	1	11	0	0
Other	2	2	0	0
<b>Degree</b>				
BA	126	77	17	25
MA	32	12	2	1
PhD	1	0	0	0
<b>TFA Status</b>				
Yes	0	32	0	2
No	159	57	19	24

Note: N = 293

From this population, teachers for whom no matches were found were removed from the analysis. The initial matching procedures produced 116 pairs of teachers (N=232 individual teachers) out of the total teacher pool of 293, thereby using 79% of the original dataset. We undertook analyses of the quality of the matching procedures that were used. These are reported on, below, and each analysis is based either on the entire sample drawn (N = 293), or on the pairs that were created by the matching procedure that was used (N = 232).

However, the process of matching produced 28 mismatched pairs consisting of teachers who did not share class score data for the same test administrations. This occurred, for example, when one teacher in a pair had scores for 1998 and the other teacher had scores for 1999. Additional matches were then made based on all of the above matching rules, but eliminating cross year matches. This finally resulted in N=109 matched pairs, using 74% of the original data set. **The appendix to this study contains descriptive information about the 109 pairs of teachers who comprise the sample for this study. The data are also available as a Microsoft Excell spreadsheet.**

Matching in primary schools across grades 3-8, however, created a problem inherent in archival studies such as this one. To design this study sensibly we needed teachers of self-contained classes. If departmentalization (more than one teacher working with the class) were occurring we would have problems inferring a teacher's affect on student learning. But we have no knowledge of what went on in every school at these upper grade levels. We were told, however, that these schools used little departmentalization, and that the teachers for whom we had files were the classroom teachers of record for the district and the state. The 218 teachers in our sample were, therefore, the teachers designated by administrators as those responsible for their student 's achievements on the SAT 9 tests. Since these were the responsible teachers we included all the matches from grades 3 to 8. Nevertheless, because we worried about the issue of departmentalization in the upper two grade levels, we ran separate analyses. One set of analyses was done with the full sample of 109 pairs of certified and under-certified teachers, and another set of analyses was done with a reduced sample, eliminating all the pairs of teachers in the 7th and 8th grade. The [appendix](#) to this report describes all 109 pairs of teachers by grade level and thus identifies which pairs were eliminated from the second analysis. In the

discussion that follows, when the second analysis using the smaller sample of pairs of teachers from grades 3-6 produces results different from the analysis of the entire sample, we will note these differences.

## **Instruments**

The Stanford Achievement Test, Ninth Edition (SAT 9), a nationally norm referenced standardized test is used by all districts in the state of Arizona and was used, as well, in this study. The test assesses student achievement in reading, math and language arts (Harcourt Brace, 1997). The SAT 9 is believed by the State Department of Education to relate to Arizona academic standards, which teachers use as a guide to instruction. It is claimed that The SAT 9 tests between 70-80% of the material outlined in the state's academic standards (ADE personal communication, 2001). This relationship is strongest for the reading and mathematics subtests, and in grades 2 through 8. The language subtest of the SAT 9 is not as well related to state standards because it does not require a writing sample of students, an ability that is promoted in the standards. The analysis of the State Department of Education suggests that the SAT 9 is a reasonable indicator of student achievement, perhaps more for reading and mathematics than for language. Furthermore, in Arizona the test is often used as an indicator of teacher and school effectiveness.

*Scoring.* Once teachers were matched to one another their Stanford Achievement Test, Ninth Edition (SAT 9) scores were aggregated at the class level. For each pair of teachers their mean National Normal Curve Equivalent (NCE) scores for SAT 9 reading, math and language were analyzed. The NCE scores are a type of normalized standard score resulting from the division of the normal curve into 99 equal units. This score is traditionally used for research purposes, enabling researchers to interpret differences in NCEs more readily because of the equal-interval nature of the NCE scores. Differences between NCEs obtained by different groups have the same meaning regardless of what part of the scale is referenced. For the purposes of this study individual student scores were not collected and thus cannot be reported.

## **Analyses of Matching Procedures**

The credibility of the matching procedure for pairing uncertified and certified teachers is important in the interpretation of the results of this study. Therefore, we undertook some analyses to explore that issue. We began by looking at the similarity of the SAT 9 test scores in each school and district to determine their comparability to each other, a check on the level of "sameness" of each school and district to one another. These analyses were conducted using NCE scores on the mathematics, reading and language sub-tests of the SAT 9, for both the 1998-1999 and 1999-2000 data sets. Alpha levels were set at  $p = .05$ .

*Matching Analysis 1.* To answer the question about whether students' test scores on the SAT 9 are different as a function of which school they attend, a one-way analysis of variance (ANOVA) was conducted on the entire data set ( $N=293$ ) to evaluate the relationship between the school assignments and student achievement scores. The independent variable was teachers' school assignment; the dependent variable was mean student achievement scores of these teachers as measured in Normal Curve Equivalents (NCE) for reading, mathematics and language in 1998-1999 and 1999-2000. The ANOVA was not significant with the exception of mathematics scores in the 1999 sample. These results are provided in Table 2. The results indicate that, overall, the schools from whom teachers in the sample came showed no statistically significant differences from each other in terms of mean NCE scores on the SAT 9,

except for the special case of mathematics scores in the 1999 sample,  $F(41,190) = 1.65, p = .01$ . It is unlikely that there are inherent differences in the schools that could bias the findings of subsequent analyses.

**Table 2**  
**ANOVA Results for School Sameness**

SAT 9 Subtest and Year		Sum of Squares	df	Mean Square	F	Sig.
Reading 1998	Between Groups	2853.47	36	79.26	1.00	0.487
	Within Groups	7389.23	93	79.45		
	Total	10242.70	129			
Math 1998	Between Groups	2801.62	36	77.82	1.07	0.385
	Within Groups	6753.20	93	72.62		
	Total	9554.82	129			
Language 1998	Between Groups	3887.73	36	107.99	1.20	0.237
	Within Groups	8341.16	93	89.69		
	Total	12228.89	129			
Reading 1999	Between Groups	4621.26	41	112.71	1.42	0.060
	Within Groups	15037.23	190	79.14		
	Total	19658.49	231			
Math 1999	Between Groups	6176.95	41	150.66	1.65	0.014*
	Within Groups	17398.93	190	91.57		
	Total	23575.88	231			
Language 1999	Between Groups	4831.13	41	117.83	1.37	0.083
	Within Groups	16363.21	190	86.12		
	Total	21194.3404	231			

Note. \* Indicates significance  $p = .05$

To determine df for 1998-1999 sample:

BG= each school with cases  $N = 37$ ,  $df = 37 - 1 = 36$

WG= each case (130) – total groups (37)  $df = 130 - 37 = 93$

Total  $df = N - 1$ ,  $130 - 1 = 129$

To determine df for 1999-2000 sample:

BG = each school with cases  $N = 42$ ,  $df = 42 - 1 = 41$

WG= each case (232) – total groups (42)  $df = 232 - 42 = 190$

Total  $df = N - 1$ ,  $232 - 1 = 231$

*Matching Analysis 2.* To answer the question about whether the test scores of students whose teachers might be paired in later analyses differed as a function of which district they attend, a one-way analysis of variance (ANOVA) was conducted on the entire data set ( $N = 293$ ) to evaluate the relationship between the district assignments and student achievement scores. The independent variable was teachers' district assignment; the dependent variable was the classroom mean student achievement scores of these teachers as measured in Normal Curve

Equivalents (NCE) for reading, mathematics and language in 1998-1999 and 1999-2000. The results indicate significant differences between the mean NCE scores for all subtests and for both years of data. For the 1998-1999 data set the ANOVA was significant,  $F(4,126) = 3.20$ ,  $p = .02$  for reading,  $F(4,126) = 2.81$ ,  $p = .03$  for mathematics, and  $F(4,126) = 3.38$ ,  $p = .01$  for language. For the 1999-2000 data set,  $F(4, 228) = 8.19$ ,  $p = .01$  for reading,  $F(4, 228) = 8.75$ ,  $p = .01$  for mathematics, and  $F(4, 228) = 6.93$ ,  $p = .01$  for language. These results are provided below in Table 3.

**Table 3**  
**ANOVA Results for District Sameness**

SAT 9 Subtest and Year		Sum of Squares	df	Mean Square	F	Sig.
Reading 1998	Between Groups	957.57	4	239.39	3.20	0.015*
	Within Groups	9413.53	126	74.71		
	Total	10371.10	130			
Math 1998	Between Groups	799.17	4	199.79	2.81	0.028*
	Within Groups	8964.88	126	71.15		
	Total	9764.05	130			
Language 1998	Between Groups	1183.68	4	295.92	3.38	0.011*
	Within Groups	11045.56	126	87.66		
	Total	12229.24	130			
Reading 1999	Between Groups	2471.10	4	617.78	8.19	.000*
	Within Groups	17200.50	228	75.44		
	Total	19671.60	232			
Math 1999	Between Groups	3144.95	4	786.24	8.75	.000*
	Within Groups	20486.51	228	89.85		
	Total	23631.46	232			
Language 1999	Between Groups	2297.24	4	574.31	6.93	.000*
	Within Groups	18902.61	228	82.91		
	Total	21199.85	232			

Note. \* Indicates significance  $p = .05$

To determine df for 1998-1999 sample:

BG= each district  $N = 5$ ,  $df = 5 - 1 = 4$

WG= each case (131) – total groups (5)  $df = 131 - 5 = 126$

Total  $df = N - 1$ ,  $131 - 1 = 130$

To determine df for 1999-2000 sample:

BG= each district  $N = 5$ ,  $df = 5 - 1 = 4$

WG= each case (233) – total groups (5)  $df = 233 - 5 = 228$

Total  $df = N - 1$ ,  $233 - 1 = 232$

The results of this ANOVA indicate that at the district level the mean student NCE scores were statistically different from one another for the teachers who comprise the population from

which our sample would be analyzed. This suggests that the procedures we used to match teachers across districts were not faultless. But analysis of the mean student scores across districts suggests that only one district may have been an outlier, with slightly higher SAT 9 scores than the others. Since only 38% of all teachers had to be matched with teachers from another district, it is likely, therefore, that only a small percent of those matches could have been problematic, totaling less than 10% of all the matches that were made. In addition, the matching of teachers across district lines was based on multiple measures of district sameness; NCE scores provide only one such measure. Because of that, we believe that the matching of teachers across district lines can still be defended as a reasonable way to obtain a sample for analysis of the student achievement of certified and uncertified teachers.

The discrepancy between the results for the ANOVA on the participating schools and the participating districts is curious and remains an issue to be resolved. This is, of course, one of the reasons that hierarchical designs have become necessary in the analysis of classroom, school and district level data. But for the purposes of this study, it is not clear that this discrepancy would cause any systematic bias in the data analyses to follow.

*Matching Analysis 3.* After the 109 pairs of matched teachers were identified we then inquired whether the average SAT 9 scores of certified teachers differed as a function of whether they were matched with teachers within their same school district or with teachers from another participating school district. A one-way analysis of variance (ANOVA) was used in this analysis. The independent variable was teachers' district assignment, either within or between school district. The dependent variable was the student achievement scores of these teachers as measured in Normal Curve Equivalents (NCE) for reading, mathematics and language in 1998-1999 and 1999-2000. ANOVA results indicate no significant differences between the mean NCE scores for the certified teachers matched within the same district as compared to certified teachers matched between districts. These results are provided below in Table 4.

**Table 4**  
**ANOVA Results for Certified Teacher Matches**  
**Within and Between Districts**

SAT 9 Subtest and Year		Sum of Squares	df	Mean Square	F	Sig.
Reading 1998	Between Groups	66.66	1	66.66	0.84	0.36
	Within Groups	4195.26	53	79.16		
	Total	4261.92	54			
Math 1998	Between Groups	56.29	1	56.29	0.76	0.39
	Within Groups	3932.17	53	74.19		
	Total	3988.46	54			
Language 1998	Between Groups	8.88	1	8.88	0.11	0.74
	Within Groups	4251.57	53	80.22		
	Total	4260.45	54			
Reading 1999	Between Groups	30.45	1	30.45	0.36	0.55
	Within Groups	7566.44	89	85.02		

	Total	7596.89	90			
Math 1999	Between Groups	4.68	1	4.68	0.05	0.82
	Within Groups	8336.88	89	93.67		
	Total	8341.56	90			
Language 1999	Between Groups	23.20	1	23.20	0.33	0.57
	Within Groups	6279.82	89	70.56		
	Total	6303.02	90			

To determine df for 1998-1999 sample:

BG= match type N= 2, df= 2-1= 1

WG= each case (55) – total groups (2) df= 55-2=53

Total df= N-1, 55-1= 54

To determine df for 1999-2000 sample:

BG= match type N= 2, df= 2-1= 1

WG= each case (91) – total groups (2) df= 91-2=89

Total df= N-1, 91-1= 90

*Matching Analysis 4.* After the 109 pairs of matched teachers were identified we then inquired whether the average SAT 9 scores of under-certified (emergency, temporary or provisional certified) teachers' differed as a function of whether they are matched within the same school district or with teachers from another participating school district. A one-way analysis of variance (ANOVA) was conducted. The independent variable was teachers' district assignment, either within or between school district. The dependent variable was the student achievement scores of these teachers, as measured in Normal Curve Equivalents (NCE) for reading, mathematics and language in 1998-1999 and 1999-2000. In general, the ANOVA results indicate no significant differences between the scores of under-certified teachers matched within district as compared to under-certified teachers matched between districts. But that was not true for all tests. Significant differences in NCE scores were found for mathematics in the 1999-2000 data set. For mathematics,  $F(1, 93) = 8.08, p = .01$ . The exclusion of 7th and 8th grade teachers yielded similar results;  $F(1, 67) = 4.12, p = .047$ . These results are provided below in Table 5.

**Table 5**  
**ANOVA Results for Under-Certified Teacher Matches**  
**Within and Between Districts**

SAT 9 Subtest and Year		Sum of Squares	df	Mean Square	F	Sig.
Reading 1998	Between Groups	0.93	1	0.93	0.02	0.90
	Within Groups	2214.40	35	63.27		
	Total	2215.33	36			
Math 1998	Between Groups	19.52	1	19.52	0.45	0.51
	Within Groups	1503.55	35	42.96		
	Total	1523.07	36			
Language 1998	Between Groups	14.09	1	14.09	0.17	0.68

	Within Groups	2936.85	35	83.91		
	Total	2950.94	36			
Reading 1999	Between Groups	69.52	1	69.52	0.82	0.37
	Within Groups	7849.48	93	84.40		
	Total	7919.00	94			
Math 1999	Between Groups	688.59	1	688.59	8.08	0.01*
	Within Groups	7928.94	93	85.26		
	Total	8617.53	94			
Language 1999	Between Groups	246.32	1	246.32	3.27	0.07
	Within Groups	6856.38	91	75.35		
	Total	7102.71	92			

Note. \* Indicates significance  $p = .05$

To determine df for 1998-1999 sample:

BG= match type  $N = 2$ ,  $df = 2 - 1 = 1$

WG= each case (37) – total groups (2)  $df = 37 - 2 = 35$

Total  $df = N - 1$ ,  $37 - 1 = 36$

To determine df for 1999-2000 sample:

BG= match type  $N = 2$ ,  $df = 2 - 1 = 1$

WG= each case (95) – total groups (2)  $df = 95 - 2 = 93$

Total  $df = N - 1$ ,  $95 - 1 = 94$

For 1999-2000 Language: Total cases = 93

It has been argued, above, that the matching procedures used in this study were sensible. The four statistical analyses intended to evaluate the matching procedures provide evidence that they were not perfect, but that evidence does not lead to the conclusion that the approach taken in this study was unreasonable or would lead to faulty conclusions. The matching of the pairs of teachers, one certified with one under-certified, within and across district lines, took place before the SAT scores of the teachers in each pair were scrutinized. Thus the matching procedures appear to be unbiased with regard to the research questions that are of interest. The results of the analyses appropriate to these research questions are considered next.

## Results

We first chose to look at whether the three kinds of under-certified teachers differed among themselves. We believed that subsequent analyses would be simpler if the SAT 9 NCE scores of the students of these three groups of teachers were not statistically different from each other. If that were the case, we could treat the three sub-groups of under-certified teachers as a single group.

A one-way analysis of variance (ANOVA) was conducted in which the independent variable was teachers' certification, while the dependent variable was the student achievement scores of these teachers as measured in Normal Curve Equivalents (NCE) for reading, mathematics and language in 1998-1999 and 1999-2000. Results indicate that the NCE scores for all under-certified teachers (emergency, temporary and provisional) were not statistically different from one another. These results are provided below in Table 6. Because of this finding we subsequently treated all uncertified teachers as a homogeneous group. In all subsequent

analyses we will compare certified teachers to the total group of under-certified teachers (emergency, temporary and provisional).

**Table 6**  
**ANOVA Results for Emergency, Temporary & Provisional Certified Teachers**

SAT 9 Subtest and Year		Sum of Squares	df	Mean Square	F	Sig.
Reading 1998	Between Groups	16.91	2	8.46	0.13	0.88
	Within Groups	2198.42	34	64.66		
	Total	2215.33	36			
Math 1998	Between Groups	68.48	2	34.24	0.80	0.46
	Within Groups	1454.59	34	42.78		
	Total	1523.07	36			
Language 1998	Between Groups	176.61	2	88.31	1.08	0.35
	Within Groups	2774.33	34	81.60		
	Total	2950.94	36			
Reading 1999	Between Groups	116.51	2	58.25	0.69	0.51
	Within Groups	7802.49	92	84.81		
	Total	7919.00	94			
Math 1999	Between Groups	131.56	2	65.78	0.71	0.49
	Within Groups	8485.97	92	92.24		
	Total	8617.53	94			
Language 1999	Between Groups	120.69	2	60.35	0.78	0.46
	Within Groups	6982.02	90	77.58		
	Total	7102.71	92			

To determine df for 1998-1999 sample:

BG= match type N= 3, df= 3-1= 2

WG= each case (37) – total groups (3) df= 37-3=34

Total df= N-1, 37-1= 36

To determine df for 1999-2000 sample:

BG= match type N= 3, df= 3-1= 2

WG= each case (95) – total groups (3) df= 95-3=92

Total df= N-1, 95-1= 94

For 1999-2000 Language: Total cases= 93

To answer the key question in this study, whether students taught by certified teachers outscore students taught by under-certified teachers, a correlated *t*-test was conducted to evaluate the difference in student achievement scores of the certified and under-certified teachers. The results indicate that for 1998-1999, students taught by certified teachers outperformed students taught by under-certified teachers. More specifically, the reading scores of the students of certified teachers were significantly higher ( $M=36.52$ ,  $SD= 9.59$ ) than were the reading scores

obtained by the students of under-certified teachers ( $M=30.67$ ,  $SD= 8.02$ ),  $t(27)= 2.36$ ,  $p= .01$ . In this same year, on the language test, the scores of the students taught by certified teachers were significantly higher ( $M=34.33$ ,  $SD= 9.17$ ) than were the scores of the students taught by under-certified teachers ( $M=29.89$ ,  $SD= 9.82$ ),  $t(27)= 1.81$ ,  $p= .04$ . While the difference between the certified and the under-certified teachers on the mathematics test were not found to be significant, the results were in the same direction as they were for the reading and language tests. Students taught by certified teachers scored higher ( $M=38.80$ ,  $SD= 8.77$ ) than did the students taught by under-certified teachers ( $M=35.82$ ,  $SD= 7.32$ ).

Results for 1999-2000 replicated the results of the data from 1998-1999. Students taught by certified teachers significantly outperformed students taught by under-certified teachers on every test. In reading, the scores of the students of certified teachers were significantly higher ( $M=35.62$ ,  $SD= 9.31$ ) than were the scores of students instructed by under-certified teachers ( $M=32.48$ ,  $SD= 9.43$ ),  $t(86)= 2.43$ ,  $p= .01$ . In mathematics, the scores of students of certified teachers were significantly higher ( $M=39.75$ ,  $SD= 9.52$ ) than were the scores obtained by students of under-certified teachers ( $M=35.22$ ,  $SD=9.77$ ),  $t(86)= 2.95$   $p= .001$ . And in the area of language, the scores of the students of certified teachers were significantly higher ( $M=35.60$ ,  $SD= 8.57$ ) than were the scores of the students instructed by under-certified teachers ( $M=33.47$ ,  $SD= 8.90$ ),  $t(84)= 1.71$ ,  $p= .05$ . These results are provided below in Table 7. The exclusion of 7th and 8th grade teachers yielded similar, and more dramatic results; the average difference between these two groups increased. Moreover, in this analysis of only grades 3-6, all subtests across both years were found to be significantly different.

**Table 7**  
**Correlated t- test Results Comparing**  
**Certified and Under-Certified Teachers**

SAT 9 Subtest and Year	Mean of Differences	St. Dev.	St. Error of Mean	95% Conf. Int. Lower Limit	95% Conf. Int. Upper Limit	t	df	Sig.
Reading 1998	5.85	13.11	2.48	0.77	10.93	2.36	27	0.01*
Math 1998	2.97	11.44	2.16	-1.47	7.41	1.37	27	0.09
Language 1998	4.44	13.01	2.46	-0.60	9.49	1.81	27	0.04*
Reading 1999	3.14	12.07	1.29	0.57	5.71	2.43	86	0.01*
Math 1999	4.53	14.31	1.53	1.48	7.58	2.95	86	0.00*
Language 1999	2.13	11.49	1.25	-.35	4.61	1.71	84	0.05*

*Note.* \* Indicates significance  $p= .05$   
 To determine df for 1998-1999 sample:  
 Number of matches  $N=28$   
 Total  $df= N-1$ ,  $28-1= 27$   
 To determine df for 1999-2000 sample:  
 Number of matches  $N=87$   
 Total  $df= N-1$ ,  $87-1= 86$   
 For 1999-2000 Language: Total cases= 85

The NCE scale provides a metric for evaluating the differences between certified teachers and under-certified teachers. Students taught by certified teachers outscored their counterparts by 6 NCE points in reading, 3 NCE points in mathematics and nearly 5 NCE points in language in

1998-1999. The results are similar for 1999-2000. Students taught by certified teachers outscored their counterparts by 3 NCE points in reading, 5 NCE points in mathematics, and 2NCE points in language. Expressed as a proportion of the standard deviation of the NCE scale represented as an effect size, these differences range across the two years from .14 to .28 in reading, .14 to .24 in mathematics and .09 to .19 in language. These results are provided below in Table 8. The exclusion of 7th and 8th grade teachers from this analysis yielded similar, yet more dramatic results in terms of effect size. Across the two years the range of the effect sizes were from .19 to .38 in reading, .24 to .28 in mathematics and .14 to .33 in language. For ease of discussion it is appropriate to choose a summary statistic to represent these data. A reasonable way to do that is to conclude that the average ES across all sub-tests of the SAT 9, across both years of testing, and across analyses, is around .20. Because of the relationship between effect size (ES) and yearly progress on standardized tests (Glass, 2002), one could expect that during one academic year in the primary grades, the students of certified teachers would make approximately 2 months more academic growth than would the students of under-certified teachers. The academic year is a 10-month year so the loss of two months or 2/10ths of a year is the loss incurred by students placed with under-certified teachers. That is, students pay approximately a 20% penalty in academic growth for each year of placement with under-certified teachers.

**Table 8**  
**NCE Differences between Certified and Under-Certified Teachers & Effect Size (ES) \* Ranges for 1998-1999 and 1999-2000**

SAT Sub-Test	1998 – 1999	1999 – 2000	ES Range
Reading	6	3	.14 – .28
Math	3	5	.14 – .24
Language	4	2	.09 – .19

Note. \*Effect sizes (ES) when using normal curve equivalencies (NCE) must be calculated with a standard deviation of 21.06 NCE units.

To answer the question whether the test scores of students of teachers in the Teach for America program are different from the scores of students who studied with other under-certified teachers, a one-way analysis of variance (ANOVA) was conducted. The independent variable was TFA trained vs. all other forms of training among the under-certified teachers. The dependent variable was the student achievement scores of these teachers as measured in Normal Curve Equivalents (NCE) for reading, mathematics and language in 1998-1999 and 1999-2000. ANOVA results indicate that the NCE scores of TFA teachers were not statistically different from the NCE scores for other under-certified teachers. These results are provided below in Table 9.

**Table 9**  
**ANOVA Results Teach for America Teachers & Other Under-Certified Teachers**

SAT 9 Subtest and Year		Sum of Squares	df	Mean Square	F	Sig.
Reading 1998	Between Groups	34.00	1	34.00	0.55	0.47

	Within Groups	2181.33	35	62.32		
	Total	2215.33	36			
Math 1998	Between Groups	21.30	1	21.30	0.50	0.49
	Within Groups	1501.78	35	42.91		
	Total	1523.08	36			
Language 1998	Between Groups	261.64	1	261.64	3.41	0.07
	Within Groups	2689.30	35	76.84		
	Total	2950.94	36			
Reading 1999	Between Groups	92.74	1	92.74	1.10	0.30
	Within Groups	7826.26	93	84.15		
	Total	7919.00	94			
Math 1999	Between Groups	1.31	1	1.31	.01	.91
	Within Groups	8616.22	93	92.65		
	Total	8617.53	94			
Language 1999	Between Groups	19.94	1	19.94	0.26	0.61
	Within Groups	7082.77	91	77.83		
	Total	7102.71	92			

To determine df for 1998-1999 sample:

BG= TFA or under-certified N= 2, df= 2-1= 1

WG= each case (37) – total groups (2) df= 37-2=35

Total df= N-1, 37-1= 36

To determine df for 1999-2000 sample:

BG= TFA or under-certified N= 2, df= 2-1= 1

WG= each case (95) – total groups (2) df= 95-2=93

Total df= N-1, 95-1= 93

For 1999-2000 Language: Total cases= 93

Given the previous analyses, in which the students of certified teachers outperformed the students of under-certified teachers and the students of TFA teachers scored no different than did the students of other under-certified teachers, it may be that the next analysis is superfluous. Nevertheless, because of the intensity of the debate about the performance of TFA teachers described in the literature reviewed above, we chose to inquire whether students taught by certified teachers outperformed students taught by Teach for America (TFA) teachers.

A correlated *t*-test was used for this analysis and it indicated that for the 1999-2000 data set, in reading, the scores of the students of the certified teachers were significantly higher ( $M=35.53$ ,  $SD= 9.87$ ) than were the scores of the students of TFA teachers ( $M=30.51$ ,  $SD= 6.85$ ),  $t(21)= 1.87$ ,  $p= .04$ . In mathematics, the scores of the students of the certified teachers were significantly higher ( $M=41.15$ ,  $SD=9.21$ ) than were those obtained by students of teachers from the TFA program ( $M=35.34$ ,  $SD= 7.67$ ),  $t(21)=2.13$ ,  $p= .02$ . Finally, in language, the scores of the students in the classes of the certified teachers were significantly higher ( $M=36.43$ ,  $SD= 9.48$ ) than were the scores of the students of teachers trained by TFA ( $M=32.11$ ,  $SD= 8.71$ ),  $t(21)=1.79$ ,  $p= .04$ .

Although the same pattern of results were found for the 1998-1999 sample, the differences between the two groups were not statistically significant. We believe this occurred because of the smaller sample size in the 1998-1999 analysis. These results are provided below in Table 10. The exclusion of 7th and 8th grade teachers yielded similar, and more dramatic results; the average difference between these two groups increased. In this analysis the differences were found to be significant in both years, in all subtests, except for math in 1998-1999.

**Table 10**  
**Correlated t-test Results Comparing**  
**Certified Teachers and Teach for America Teachers**

SAT 9 Subtest and Year	Mean of Differences	St. Dev.	St. Error of Mean	95% Conf. Int. Lower Limit	95% Conf. Int. Upper Limit	t	df	Sig.
Reading 1998	4.28	10.10	3.57	-4.17	12.73	1.20	7	0.13
Math 1998	2.25	10.02	3.54	-6.13	10.62	0.63	7	0.27
Language 1998	2.57	7.15	2.53	-3.41	8.54	1.02	7	0.17
Reading 1999	5.02	12.58	2.68	-0.56	10.60	1.87	21	0.04*
Math 1999	5.81	12.81	2.73	0.13	11.49	2.13	21	0.02*
Language 1999	4.31	11.29	2.41	-0.69	9.32	1.79	21	0.04*

*Note.* \* Indicates significance  $p = .05$

To determine df for 1998-1999 sample:

Number of matches  $N=8$

Total  $df = N-1, 8-1 = 7$

To determine df for 1999-2000 sample:

Number of matches  $N=22$

Total  $df = N-1, 22-1 = 21$

The data set was also examined to gain information about the role of experience in developing teacher competency. For teachers that were in both the 1998-1999 and the 1999-2000 set of data, we had hoped to look at whether teacher experience effects student SAT 9 scores, and more particularly, whether the differences in performance between the certified teachers and the uncertified teachers was moderated as a function of the increased experience of the uncertified teachers. But the sample of teachers for whom we had data across two years was very small (six pairs across the two years), so no confident answers to these questions can be offered.

One of our analyses was a one-way within-subjects ANOVA, with the factor being experience, as measured in time from 1998-1999 to 1999-2000. The dependent variable was the student's achievement scores for these teachers as measured in Normal Curve Equivalents (NCE) for reading, mathematics and language in 1998-1999 and 1999-2000. The results indicate that there is no significant difference in NCE scores from the first year to the second year. Nevertheless, the scores for each subtest of the SAT increased from the first year to the second, indicating that teacher experience may affect the achievement test scores of their students. The means and standard deviations are provided below, in Table 11.

The scores increased from one to two NCE points in each of the three subtests, with the increase in mathematics being the greatest. The difference in the scores between the first year and second year are provided in Table 12. We also ran an ANOVA on these changes over time,

and those results are given in Table 13.

**Table 11**  
**Means and Standard Deviations for**  
**Matched Teachers with Two Years of Data**

SAT 9 Subtest and Year	Mean	Std. Deviation	N
Reading 1998-1999	36.50	14.25	12
Reading 1999-2000	37.79	7.26	12
Math 1998-1999	39.03	11.49	12
Math 1999-2000	41.07	8.38	12
Language 1998-1999	35.85	13.00	12
Language 1999-2000	37.05	7.38	12

**Table 12**  
**Difference in Mean SAT 9 Scores for**  
**Matched Teachers with Two Years of Data**

SAT 9 Subtest and Year	Mean of Differences	St. Dev.	St. Error of Mean	95% Conf. Int. Lower Limit	95% Conf. Int. Upper Limit	t	df	Sig.
Reading	-1.29	14.54	4.20	-10.52	7.95	-0.31	11	0.38
Math	-2.04	9.29	2.68	-7.95	3.86	-0.76	11	0.23
Language	-1.20	11.71	3.38	-8.64	6.24	-0.36	11	0.36

**Table 13**  
**ANOVA for Teachers with Two Years of Data**

SAT 9 Subtest	F	df	Error df	Sig.
Reading	.09	1	11	.77
Mathematics	.58	1	11	.46
Language	.13	1	11	.73

In order to evaluate whether the *differences* between certified teachers and under-certified teachers, with two years of data, remained similar, grew or decreased from the first year to the second, mean NCE scores for each group were analyzed. Results indicate that the difference between the scores of certified teachers and the scores of under-certified teachers for the 1998-1999 to 1999-2000 data set, as measured in NCE scores, decreased in reading and language, but increased in mathematics. These results are provided below, in Table 14.

**Table 14**  
**Difference between Certified and Under-Certified Teachers'**  
**NCE Scores from 1998-1999 to 1999-2000 for**

## Teachers with Two Years of Data

Academic Year	Reading	Mathematics	Language
1998-1999	12.93	7.80	9.51
1999-2000	3.18	9.47	4.02

Note. All scores favor certified teachers over under-certified teachers.

## Discussion and Conclusion

Many different values necessarily come into play when making educational policy about the qualifications that are needed to become a beginning teacher. So much is riding on the performance of these individuals, trusted with educating our nation's young. So many skills are needed to do that job well. Thus, a single empirical study of this kind cannot provide answers to complex policy questions about the relative benefits and liabilities of allowing certified and under-certified teachers to teach our young. Nevertheless, there is every reason to think that the results of this study are generalizable and worth considering when educational policies on these issues are debated.

As we understand the national situation it appears not to be very different from that in Arizona. From New York, through Chicago, and on to Los Angeles, teachers in schools that serve the poor are often under-certified, inexperienced, and may be teaching out-of-field. Teachers who serve wealthier students overwhelmingly hold regular certification, have accumulated considerably more teaching experience, and are less often required to teach out-of-field. (Darling-Hammond, 1997a, 1997b; 2001; Ingersoll, 2001; Gitomer, Latham, & Ziomek, 1999; Lankford, Loeb, & Wycoff, 2002).

This study addressed one of these factors—the effectiveness of certification on student achievement. We found what might be expected of those who choose to do complex work, namely, that those who trained longer and harder to do that work do it better. Common sense and empirical data agree. Despite our lack of understanding of how it is accomplished, and despite the extreme variability in the programs of instruction (surely masking both excellent and dreadful programs), the present research study supports the assertion that university prepared teachers are of higher quality than those prepared without an approved program of preparation (see also Evertson, 1984; Darling-Hammond, 1997a).

In this study regularly certified teachers significantly outperformed under-certified teachers with children who are most at risk of school failure and school dropout. These already low achieving children, when assigned to the classrooms of under-certified teachers made gains that were approximately 2 months less per school year on three different subtests of the SAT 9. This is about 20% less academic growth than they would have made had they been assigned to a teacher with regular state certification.

The Rowan et al. (in press) study, cited above, states that the relationship between measures of student growth and measures of teacher competency are much stronger than are the relationships found when a single years measure of achievement is used as the dependent variable, as in this study. Since the districts we studied had relatively large percentages of under-certified teachers the odds of a student getting more than one such teacher during their primary grades is high. If the magnitude of the effects on student achievement growth over

time were as high as Rowan et al. believe, then it is likely that exposure to just two under-certified teachers would result in intractable deficits in academic growth in reading, mathematics and language. Although their research methods are hard to follow, Sanders and Rivers (1996) reach similar conclusions: the effects of poor quality teachers are cumulative.

In this era of accountability for schools and students, low test performance can mean the loss of employment for teachers and administrators, while for students, such results can lead to retention in grade or denial of a high school degree. But there are school systems throughout the nation that make regular use of large numbers of under-certified teachers and thus, through their hiring practices, virtually guarantee that their students will achieve relatively low levels of performance on norm-referenced standardized tests. Students, teachers and administrators will each be made to pay for a policy that assures less than desirable outcomes from the school system.

This situation raises broad questions of policy, such as, what are the causes of, and who is accountable for, the placement of the under-certified teachers in the classrooms of our most challenging students? Who should accept responsibility for an educational policy that appears harmful and that clearly handicaps students in the lower social classes? Will the school districts that make heavy use of under-certified teachers all violate the new federal guidelines, since under-certified teachers seem *not* to be highly qualified to teach? And if these districts will not be in compliance with the new federal regulations because they cannot attract qualified teachers to their classrooms what can they do differently to receive funding and change the working conditions so that they can attract and keep qualified teachers?

Policy makers should take the results of this study seriously, perhaps also funding more research of this kind to ensure the validity of our findings. But meanwhile, on the basis of our findings, we see evidence of a harmful educational policy. We believe that those in authority need to attend to the legal and moral issues that arise from our data. It appears that we are systematically providing an inferior education to the children of the poor. They start with academic difficulties and then through the policies we adopt we handicap them 20% more per year when we assign them to classrooms staffed by under-certified teachers.

The data we have collected also inform us that there is no difference between the performance of new teachers from Teach for America and that of all other under-certified teachers. On all tests, and in both years, the certified teachers out-performed the under-certified novice teachers from Teach for America. Our results contradict claims made by TFA advocates that the enthusiasm and subject-matter knowledge, as well as a general education in a prestigious university, prepare these recruits to teach adequately in America's classrooms. The TFA teachers are no better able to teach than any other under-prepared teacher.

In general, research on Teach for America has been limited and the results are often contradictory (Darling-Hammond, 1994; 1997a; 2001; Stevens & Dial, 1993; Schorr, 1993; Kopp, 1994; Raymond, Fletcher, & Luque, 2001). Our findings do, however, directly contradict those reported by Raymond, Fletcher & Luque (2001). We find no evidence to support their claim that TFA teachers produce a positive effect on their students' achievement relative to teachers recruited in other ways. In our view, the preponderance of the available literature raises serious concerns about the TFA program. Although new TFA teachers are required to take a six-week summer training program before their school year begins, and they receive support throughout the school year from TFA personnel, the performance of their students is indistinguishable from that of student's taught by other under-certified teachers. More important for policy makers is that the level of performance of the students of the TFA

teachers was lower than that of the students taught by equally inexperienced but fully certified teachers. That is the more important finding.

TFA may be a meaningful way for young college graduates to make some money and take a few years out of the ordinary path their careers demand. But they are hurting our young, vulnerable, inner-city students. (We expect that TFA teachers are faring no better in rural communities, but our data does not address that population.) Because an overwhelmingly high percent of the TFA students also leave the profession after their two years of service, their hard earned teaching experience will never be put to use with future generations of students.

While the TFA program appears to be a failure, it is simply part of the larger pattern of failure that attends to the policy of hiring under-certified teachers. The policy of hiring under prepared teachers for the schools that serve America's poor looks like an act of class warfare, a concept that Americans find hard to accept. But states have adopted, or allowed policies to continue unchallenged, that prevent poor and rural American children from receiving the education they need for citizenship or to compete in the economy of the 21st century. These are policies to be ashamed of and abandoned. We hope that the new federal legislation will change things, for if it is taken literally, we might eventually have highly qualified teachers for all of our nations' students to learn from.

## Note

<sup>1</sup> This article is based on the first author's dissertation titled *Teacher certification does matter: The effects of certification status on student achievement*, completed Spring, 2002, in the College of Education, Arizona State University. The second author received partial funding for helping with this research from the Rockefeller Foundation, to whom we are grateful. The views expressed in this report, however, are the sole responsibility of its authors and may not reflect the views of The Rockefeller Foundation or the Arizona Department of Education.

## References

- Angus, D. L. (2001). *Professionalism and the public good: A brief history of teacher certification*. Washington, DC: Fordham Foundaton. Retrieved July 25, 2002 from <http://www.edexcellence.net/library/angus/angus.html>
- Arizona Department of Education Web Page. (2002). Available: <http://www.ade.az.gov>
- Ashton, P., & Crocker, L. (1987). Systematic study of planned variations: The essential focus of teacher education reform. *Journal of Teacher Education*, 38(1), 2-8.
- Ashton, P., Crocker, L., & Olejnik, S. (1986). *Does teacher education make a difference? A literature review and planning study*. Executive summary and technical monograph prepared for the Institute on Student Assessment and Evaluation, Florida Department of Education.
- Ballou, D. (1996). Do public schools hire the best applicants? *Quarterly Journal of Economics*, 3(2), 97-134.
- Ballou, D., & Podgursky, M. (1999). Reforming teacher preparation and licensing: What is the evidence? *Teachers College Record*. Retrieved November 1999 from <http://www.tcrecord.org/>
- Berry, B. (2001). No shortcuts to preparing good teachers. *Educational Leadership*, 58(8).

Retrieved January 2002 from <http://ascd.org/readingroom/edlead/0105/berry.html>.

Bliss, T. (1992). Alternative certification in Connecticut: Reshaping the profession. *Peabody Journal of Education*, 67(3), 35-54.

Bradshaw, L., & Hawk, P. (1996). *Teacher certification: Does it really make a difference in student achievement?* Greenville, NC: Eastern North Carolina Consortium for Assistance and Research in Education.

Buck, G. H., Polloway, E. A., & Robb, S. M. (1995). Alternative certification programs: A national survey. *Teacher Education and Special Education*, 18(1), 39-48.

Burbules, N. C., & Densmore, K. (1991). The limits of making teaching a profession. *Educational Policy*, 5(1), 44-63.

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 Northeastern Educational Research Association, Ellenville, NY.

Cochran-Smith, M. & Fries, M. K. (2001). Sticks, stones and ideology: The discourse of reform in teacher education. *Educational Researcher*, 30(8), 3-15.

Crehan, K. D., Hess, R., Lawrence, A., & Sabers, D. (2000). *Validity issues in teacher certification testing*. Paper presented at the annual meeting of Arizona Educational Research Organization, Tucson Arizona.

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

Darling-Hammond, L. (1994). Who will speak for the children? How "Teach for America" hurts urban schools and students. *Phi Delta Kappan*, 76(1), 21-34.

Darling-Hammond, L. (1997a). *The right to learn: A blueprint for creating schools that work*. San Francisco, CA: Jossey-Bass.

Darling-Hammond, L. (1997b). *Doing what matters most: Investing in quality teaching*. NY: National Commission on Teaching and America's Future, Teachers College, Columbia University. Retrieved July 27, 2002 from <http://www.nctaf.org/publications/index.html>

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1). Retrieved January 2001 from <http://epaa.asu.edu/epaa/v8n1/>

Darling-Hammond, L. (2001). *The research and rhetoric on teacher certification: A response to "Teacher Certification Reconsidered."* NY: National Commission on Teaching and America's Future, Teachers College, Columbia University Retrieved July 26, 2002 from <http://www.nctaf.org/publications/index.html>

Darling-Hammond, L. (2002). The research and rhetoric on teacher certification: A response to "Teacher Certification Reconsidered." *Educational Policy Analysis Archives*, 10(36). Retrieved September 5, 2002 from <http://epaa.asu.edu/epaa/v10n36.html>

- Darling-Hammond, L., Berry, B. & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Education Evaluation and Policy Analysis*, 23(1), 57-77.
- Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? *Journal of Teacher Education*, 53(4), 286-302.
- Eifler, K., & Potthoff, D. (1998). Nontraditional teacher education students: A synthesis of the literature. *Journal of Teacher Education*, 49(3), 187-195.
- Educational Testing Service (2002). *PRAXIS information web page*. Retrieved July 25, 2002 from <http://www.teachingandlearning.org/licensure/praxis/prxtest.html>
- Ehrenberg, R. G., & Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from High School and Beyond. *Economics of Educational Review*, 13(1), 1-17.
- Evertson, C. (1984). *The characteristics of effective preparation programs: A review of research*. Washington, DC: National Commission on Excellence in Teacher Education.
- Everston, C., Hawley, W., & Zlotnik, M. (1985). Making a difference in educational quality through teacher education. *Journal of Teacher Education*, 36(3), 2-12.
- Feistritzer, E. (1993). National overview of alternative teacher certification. *Education and Urban Society*, 26(1), 18-28.
- Feistritzer, E. (1994). Evolution of alternative teacher certification. *Educational Forum*, 58, 132-138.
- Fenstermacher, G. (1992). The place of alternative certification in the education of teachers. *Peabody Journal of Education*, 67(3), 155-185.
- Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation*, 28(2), 465-498.
- Ferguson, R. F., & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education*, 44(1), 55-63.
- Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives*, 7(9). Retrieved January 2000 from <http://epaa.asu.edu/epaa/v7n9.html>.
- Finn, C., & Kanstroom, M. (2000). Improving, empowering, dismantling. *The Public Interest*. Manhattan Institute: Authors. Retrieved December 2001 from [http://www.manhattan-institute.org/html/\\_pi-\\_improving.html](http://www.manhattan-institute.org/html/_pi-_improving.html).
- Galambos, E. C. (1984). *Testing teachers for certification and recertification*. Washington, DC: National Commission on Excellence in Teacher Education.
- Gitomer, D., Latham, A. S., & Ziomek, R. (1999). *The academic quality of prospective teachers: The impact of admissions and licensure testing*. Princeton, NJ: Teaching and Learning Division, Educational Testing Service. Retrieved July 27, 2002, from

<http://www.ets.org/teachingandlearning/rschnews.html#impact>

Glass, G.V (1978). Standards and criteria. *Journal of Educational Measurement*, 15(4), 237-61.

Glass, G.V (2002). Teacher characteristics. In A. Molnar (Ed.), *School Reform Proposals: the Research Evidence* (Chapter 8). Retrieved March 2002 from <http://www.asu.edu/educ/eps/Reports/epru/EPRU2002-101/epru-2002-101.html>.

Go, K. (2002, July 31). 1 out of 6 teachers in Arizona lacks certification. *Arizona Republic*. Retrieved August 2 from <http://www.arizonarepublic.com/news/articles/0731emergency31.html>

Goldhaber, D. D., & Brewer, D. J. (1996). *Evaluating the effect of teacher degree level*. Rockville, MD: Westat, Inc.

Goldhaber, D. D. & Brewer, D. J. (2000). Does certification matter? High school teacher certification status and student achievement. *Education Evaluation and Policy Analysis*, 22(2), 129-145.

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

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2001, November). *Why public schools lose teachers* (Working Paper 8599). Washington, DC: National Bureau of Economic Research.

Harcourt Brace Educational Measurement. (1997). *Stanford Achievement Test, Ninth Edition Spring norms book*. San Antonio, TX: Author.

Harcourt Brace Educational Measurement. (2001). *Stanford Achievement Test, Ninth Edition Administrator's interpretive guide: Arizona student achievement program*. San Antonio, TX: Author.

Hardy, L. (1998). A good teacher is hard to find. *American School Board Journal*, 185(9), 20-23.

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

Hawkins, E. F., Stancavage, F. B., & Dorsey, J. A. (1998). *School policies affecting instruction in mathematics*. Washington, D.C.: National Center for Education Statistics.

Hawley, W. (1992). The theory and practice of alternative certification: Implications for the improvement of teaching. *Peabody Journal of Education*, 67(3), 3-34.

Haycock, K. (2001). Closing the achievement gap. *Educational Leadership*, 58(6). Retrieved February, 2002, from <http://www.ascd.org/readingroom/edlead/0103/haycock.html>

Hernandez, R. (1998, December 4). Suit says minority students receive an inferior education. *The New York Times*. Retrieved, May 2002, from <http://www.nytimes.com/>

Higher Education Reauthorization Act. (1998). Pub. L. No. 105-244, 20 U.S.C. § 1021-1030.

Hirsh, E., Koppich, J. E., & Knapp, M. S. (2001). *Revisiting what states are doing to improve*

*the quality of teaching: An update on patterns and trends.* Seattle: University of Washington, Center for the Study of Teaching and Policy.

Holmes Group. (1986). *Tomorrow's teachers: A report of the Holmes Group.* East Lansing, MI: Author.

Houston, R. W., Marshall, F., & McDavid, T. (1993). Problems of traditionally prepared and alternatively certified first-year teachers. *Education and Urban Society*, 26(1), 78-89.

Howey, K., & Zimpher, N. (1994). Nontraditional contexts for learning to teach. *The Educational Forum*, 58, 156-161.

Hundreds of teachers in city lack qualifications. (2002, January 31) *Philadelphia Inquirer*. Retrieved July 26, 2002 from <http://www.edpolicy.org/rgn.iss/2002/ne/week5.htm#article2>

Ingersoll, R. (1996). *Teacher quality and inequality.* Paper presented at the annual meeting of the American Statistical Association, Chicago, IL.

Ingersoll, R. (1997). Teacher turnover and teacher quality: The recurring myth of teacher shortages. *Teachers College Record*, 91(1), 41-44.

Ingersoll, R. (1999a). Invited commentary: Understanding the problem of teacher quality in American schools. *Education Statistics Quarterly*: National Center for Education Statistics. Retrieved October 2001 from <http://www.nces.ed.gov/pubs99/quarterlyapr/2-feature/2-esq11-c.html>.

Ingersoll, R. (1999b). The problem of under-qualified teachers in American secondary schools. *Educational Researcher*, 28(2), 26-37.

Ingersoll, R. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal*, 38 (3), 499-534.

Jaeger, R. (1990). Establishing standards for teacher certification tests. *Educational Measurement: Issues and Practice*, 9(4), 15-20.

Jelmsberg, J. (1996). College-based teacher education versus state-sponsored alternative programs. *Journal of Teacher Education*, 47(1), 60-66.

Kain, J. F. (1998). *Impact of individual teachers and peers on individual student achievement.* Paper prepared for the annual meeting of the Association for Public Policy Analysis and Management, New York, NY.

Kanstroom, M., & Finn, C. (Eds.)(1999). *Better teachers better schools.* Washington, DC: Education Leaders Council, Thomas B. Fordham Foundation. Retrieved July 26, 2002, from <http://www.edexcellence.net/better/tchrs/btbs.htm>

Kearns, J. (1984). State competency testing for teacher certification. *Journal of Teacher Education*, 35(2), 58-61.

Keltner, D. (1994). Troops to teachers: Alternative certification and the military. *The Educational Forum*, 58, 182-186.

Kennedy, M. (1991). Policy issues in teacher education. *Phi Delta Kappan*, 72(9), 659-665.

Kopp, W. (1994). Teach for America: Moving beyond the debate. *The Educational Forum*, 58, 187-192.

Kwiatkowski, M. K. (1999). Debating alternative teacher certification: A trial by achievement. In M. Kanstroom & C. Finn (Eds.)(1999). *Better teachers better schools* (pp. 215-237). Washington, DC: Education Leaders Council, Thomas B. Fordham Foundation. Retrieved July 26, 2002, from <http://www.edexcellence.net/better/tchrs/btbs.htm>

Labaree, D. (1992). Power, knowledge, and the rationalization of teaching: A genealogy of the movement to professionalize teaching. *Harvard Educational Review*, 62(2), 123-154.

Lankford, H., Loeb, S., & Wycoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1), 37-62.

Lopez, O. S. (1995). *Classroom diversification: An alternative paradigm for research in educational productivity*. Unpublished doctoral dissertation, University of Texas, Austin.

Mathews, J. (2002, July 18). Teacher training: Too much or not enough? *Washington Post*. Retrieved July 20 from <http://www.washingtonpost.com/wp-dyn/articles/A41591-2002jul18.html>

MacDonald, R., Manning, M. L., & Gable, R. (1994). An alternative certification program: Career transition. *Action in Teacher Education*, 16(2), 19-27.

McDiarmid, G.W., & Wilson, S. (1991). An exploration of the subject matter knowledge of alternative route teachers: Can we assume they know their subject? *Journal of Teacher Education*, 42(2), 93-103.

McKibbin, M. (1988). Alternative teacher certification programs. *Educational Leadership*, 46(3), 32-35.

McKibbin, M., & Ray, L. (1994). A guide for alternative certification program improvement. *The Educational Forum*, 58, 201-208.

Miles-Nixon, C., & Holloway, P. (1997). Alternative certification: The good, the bad and the ugly. *Delta Kappa Gamma Bulletin*, 63(4), 40-47.

Miller, J. W., McKenna, M. C., & McKenna, B. A. (1998). A comparison of alternatively and traditionally prepared teachers. *Journal of Teacher Education*, 49(3), 165-176.

Monk, D. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Educational Review*, 12(2), 125-142.

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

Murnane, R. J. (1983). Understanding the sources of teaching competence: Choices, skills and the limits of training. *Teachers College Record*, 83(3), 564-569.

- Murnane, R. J. (1991). The case for performance based licensing. *Phi Delta Kappan*, 73(2), 137-142.
- Mitchell, K. J., Robinson, D. Z., Plake, B. S., & Knowles K. T. (Eds). (2001). *Testing teacher candidates: The role of licensure tests in improving teacher quality*. National Research Council, Washington, D.C: National Academy Press.
- Natriello, G., & Zumwalt, K. (1993). New teachers for urban schools? The contribution of the provisional teacher program in New Jersey. *Education and Urban Society*, 26(1), 49-62.
- No Child Left Behind (NCLB). (2001). Conference report to accompany H.R. Rep. No. 107-334, 107th Cong., 1st sess. (2001).
- Olson, L. (2000). Finding and keeping competent teachers. *Education Week*, 19(18), 12-18.
- Pearce, K. (2000, August 18). Substitutes can teach without college degree. *The Arizona Republic*, pp. B1, B2.
- Peck, H. (1989). *The effect of certification status on the performance of mathematics teachers: A pilot study*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Pratte, R. & Rury, J. L. (1991). Teachers, professionalism, and craft. *Teachers College Record*, 93(1), 59-72.
- Pyburn, K. (1990). Legal challenges to licensing examinations. *Educational Measurement: Issues and Practice*, 9(4), 5-6, 14.
- Quality counts 2000: Who should teach? The states decide. (2000). *Education Week*, 19(18).
- Quirk, T. J., Witten, B. J., & Weinberg, S. F. (1973). Review of studies of the concurrent and predictive validity of the National Teacher Examinations. *Review of Educational Research*, 43(1), 89-113.
- Raymond, M., Fletcher, S., & Luque, J. (2001) *Teach For America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Palo Alto, CA: CREDO, The Hoover Institution, Stanford University. Retrieved September 1, 2002 from <http://www.rochester.edu/credo>.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2000). *Teachers, schools and academic achievement*. Working Paper. Retrieved March 2001 from <http://www.utdallas.edu/research/greenctr/Papers>.
- Roth, R. (1994). The university can't train teachers? Transformation of a profession. *Journal of Teacher Education*, 45(4), 261-268.
- Rowan, B., Correnti, R., & Miller, R. J. (in press). What large-scale, survey research tells us about teacher effects on student achievement: insights from the *Prospects* study of elementary schools. *Teachers College Record*.
- Sanders, W. L. (1997). *Teacher and classroom context effects on student achievement*:

*Implications for teacher evaluation.* Retrieved March 2001 from <http://www.mdk12.org/practices/ensure/tva>.

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

Schaeffer, B. (1996). *Standardized tests and teacher competence.* Fair Test. Retrieved January 2000 from <http://www.fairtest.org/empl/ttcomp.html>.

Schorr, J. (1993). Class Action. *Phi Delta Kappan*, 7(4), 315-318.

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.

Scores needed to pass Minnesota teacher tests among lowest in U. S. (2002, July 21). *Minneapolis Star Tribune*. Retrieved July 26, 2002 from <http://www.startribune.com/stories/462/3090436.htm>

Shimberg, B. (1990). Social considerations in the validation of licensing and certification exams. *Educational Measurement: Issues and Practice*, 9(4), 11-14.

Smith, I. L., & Hambleton, R. K. (1990). Content validity studies of licensing examinations. *Educational Measurement: Issues and Practice*, 9(4), 7-10.

Smith, J. (1991). The alternate route: Flaws in the New Jersey plan. *Educational Leadership*, 49, 32-36.

Soar, R. S., Medley, D. M., & Coker, H. (1983). Teacher evaluation: A critique of currently used methods. *Phi Delta Kappan*, 64(4), 239-246.

Sowell, T. (2002, July 31). Unsettling school data. *Washington Times*. Retrieved on August 2, 2002, from <http://www.washingtontimes.com/commentary/20020731-87081079.htm>

Stevens, C. J., & Dial, M. (1993). A qualitative study of alternatively certified teachers. *Education and Urban Society*, 26(1), 63-77.

Stoddart, T. (1992). Los Angeles Unified School District intern program: Recruiting and preparing teachers for the urban context. *Peabody Journal of Education*, 26, 29-48.

Stoddart, T. (1993). Who is prepared to teach in urban schools? *Education and Urban Society*, 26(1), 29-48.

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

Students of certified teachers more likely to pass skills test (1999, April 12). *Abilene Reporter*. Retrieved on July 26, 2002 from <http://www.texnews.com/1998/1999/texas/test0412.html>

Summers, A., & Wolfe, B. (1975). *Which school resources help learning? Efficiency and equity in Philadelphia public schools.* Business Review. Federal Bank of Philadelphia, PA.

U.S. Department of Education, National Center for Education Statistics. (2001). *Education finance statistics center peer search detail*. Washington, DC: Author. Retrieved August 2001 from <http://www.nces.ed.gov/edfin/search/>.

U.S. Department of Education, National Center for Education Statistics. (1999). (Lewis, L., Parsad, B., Carey, N., Bartfai, N., Farris, E., & Smerdon, B.) *Teacher quality: A report on the preparation of public school teachers*. Washington, DC: Author.

Wayne, A. (2000). Teacher supply and demand: Surprises from primary research. *Educational Policy Analysis Archives*, 8(47). Retrieved December 2000 from <http://epaa.asu.edu/epaa/v8n47.html>.

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

Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2002). Teacher preparation research: An insider's view from the outside. *Journal of Teacher Education*, 53(3), 190-204.

Wise, A. (1994a). The coming revolution in teacher licensure: Redefining teacher preparation. *Action in Teacher Education*, 16(2), 1-13.

Wise, A. (1994b). Choosing between professionalism and amateurism. *The Educational Forum*, 58, 139-146.

Wong, E. (2000, August 13). Poorest schools lack teachers and computers. *The New York Times*. Retrieved August, 2000, from <http://www.nytimes.com/>

Wyatt, E. (2000, July 14). Regents support plan to close some Roosevelt schools if standards are not met. *The New York Times*. Retrieved July, 2000 from <http://www.nytimes.com/>

Zumwalt, K. (1991). Alternative routes to teaching: Three alternative approaches. *Journal of Teacher Education*, 42(2), 83-92

## Appendix

In HTML  
In Excell Spreadsheet

### About the Authors

#### Ildiko Laczko-Kerr

Arizona Department of Education  
Phoenix, AZ

Ildiko Laczko-Kerr is an Education Research Associate in the Research and Policy Unit at the Arizona Department of Education. She received her Ph.D. in Educational Psychology in 2002

from the College of Education at Arizona State University. She holds a BA in Psychology and a MA in Educational Psychology from the University of Arizona. Her current research at the department of education involves working with state assessment data, the development and implementation of the state's achievement profiles, and research relating to Arizona's teaching force. She can be reached by email at [ilaczko@ade.az.gov](mailto:ilaczko@ade.az.gov).

**David C. Berliner**

Regents' Professor of Education  
College of Education  
Arizona State University  
Tempe, AZ 85287-2411

Email: [berliner@asu.edu](mailto:berliner@asu.edu)

David C. Berliner is Regents' Professor of Education at the College of Education of Arizona State University, in Tempe, AZ. He received his Ph.D. in 1968 from Stanford University in educational psychology, and has held positions at the University of Massachusetts, WestEd, and the University of Arizona. He has served as president of the American Educational Research Association (AERA), president of the Division of Educational Psychology of the American Psychological Association (APA), and as a Fellow of the Center for Advanced Study in the Behavioral Sciences. Berliner is a member of the National Academy of Education. His publications include *The Manufactured Crisis*, Addison-Wesley, 1995 (with B.J. Biddle) and the *Handbook of Educational Psychology*, Macmillan, 1996 (Edited with R.C. Calfee). Special awards include the Research into Practice Award and the lifetime achievement award from AERA, the E. L. Thorndike award from APA, and the 2003 Brock international award for educational achievements. His scholarly interests include research on teaching and education policy analysis.

---

Copyright 2002 by the *Education Policy Analysis Archives*

The World Wide Web address for the *Education Policy Analysis Archives* is [epaa.asu.edu](http://epaa.asu.edu)

General questions about appropriateness of topics or particular articles may be addressed to the Editor, [Gene V Glass](mailto:Gene V Glass), [glass@asu.edu](mailto:glass@asu.edu) or reach him at College of Education, Arizona State University, Tempe, AZ 85287-2411. The Commentary Editor is Casey D. Cobb: [casey.cobb@unh.edu](mailto:casey.cobb@unh.edu) .

**EPAA Editorial Board**

[Michael W. Apple](#)  
University of Wisconsin

[John Covalleskie](#)  
Northern Michigan University

[Sherman Dorn](#)  
University of South Florida

[Richard Garlikov](#)  
[hmwkhelp@scott.net](mailto:hmwkhelp@scott.net)

[Greg Camilli](#)  
Rutgers University

[Alan Davis](#)  
University of Colorado, Denver

[Mark E. Fetler](#)  
California Commission on Teacher Credentialing

[Thomas F. Green](#)  
Syracuse University

Alison I. Griffith  
York University

Ernest R. House  
University of Colorado

Craig B. Howley  
Appalachia Educational Laboratory

Daniel Kallós  
Umeå University

Thomas Mauhs-Pugh  
Green Mountain College

William McInerney  
Purdue University

Les McLean  
University of Toronto

Anne L. Pemberton  
apembert@pen.k12.va.us

Richard C. Richardson  
New York University

Dennis Sayers  
California State University—Stanislaus

Michael Scriven  
scriven@aol.com

Robert Stonehill  
U.S. Department of Education

Arlen Gullickson  
Western Michigan University

Aimee Howley  
Ohio University

William Hunter  
University of Ontario Institute of  
Technology

Benjamin Levin  
University of Manitoba

Dewayne Matthews  
Education Commission of the States

Mary McKeown-Moak  
MGT of America (Austin, TX)

Susan Bobbitt Nolen  
University of Washington

Hugh G. Petrie  
SUNY Buffalo

Anthony G. Rud Jr.  
Purdue University

Jay D. Scribner  
University of Texas at Austin

Robert E. Stake  
University of Illinois—UC

David D. Williams  
Brigham Young University

## **EPAA Spanish Language Editorial Board**

**Associate Editor for Spanish Language**  
**Roberto Rodríguez Gómez**  
**Universidad Nacional Autónoma de México**

roberto@servidor.unam.mx

Adrián Acosta (México)  
Universidad de Guadalajara  
adrianacosta@compuserve.com

Teresa Bracho (México)  
Centro de Investigación y Docencia  
Económica-CIDE  
bracho dis1.cide.mx

Ursula Casanova (U.S.A.)  
Arizona State University  
casanova@asu.edu

Erwin Epstein (U.S.A.)  
Loyola University of Chicago

J. Félix Angulo Rasco (Spain)  
Universidad de Cádiz  
felix.angulo@uca.es

Alejandro Canales (México)  
Universidad Nacional Autónoma de  
México  
canalesa@servidor.unam.mx

José Contreras Domingo  
Universitat de Barcelona  
Jose.Contreras@doe.d5.ub.es

Josué González (U.S.A.)  
Arizona State University

Eepstein@luc.edu

**Rollin Kent (México)**

Departamento de Investigación  
Educativa-DIE/CINVESTAV  
rkent@gemtel.com.mx  
kentr@data.net.mx

**Javier Mendoza Rojas (México)**

Universidad Nacional Autónoma de  
México  
javiermr@servidor.unam.mx

**Humberto Muñoz García (México)**

Universidad Nacional Autónoma de  
México  
humberto@servidor.unam.mx

**Daniel Schugurensky**

(Argentina-Canadá)  
OISE/UT, Canada  
dschugurensky@oise.utoronto.ca

**Jurjo Torres Santomé (Spain)**

Universidad de A Coruña  
jurjo@udc.es

josue@asu.edu

**María Beatriz Luce (Brazil)**

Universidad Federal de Rio Grande do  
Sul-UFRGS  
lucemb@orion.ufrgs.br

**Marcela Mollis (Argentina)**

Universidad de Buenos Aires  
mmollis@filo.uba.ar

**Angel Ignacio Pérez Gómez (Spain)**

Universidad de Málaga  
aiperez@uma.es

**Simon Schwartzman (Brazil)**

American Institutes for Resesarch–Brazil  
(AIRBrasil)  
simon@airbrasil.org.br

**Carlos Alberto Torres (U.S.A.)**

University of California, Los Angeles  
torres@gseisucla.edu