

MPR Reference No.: 6215-400

MATHEMATICA
Policy Research, Inc.

**An Evaluation of
American Board Teacher
Certification: Progress
and Plans**

May 3, 2006

*Steven Glazerman
Christina Tuttle*

Submitted to:

American Board for Certification of
Teacher Excellence
1225 19th Street, N.W., Suite 400
Washington, DC 20036
Telephone: (202) 261-2620

Project Officer:
Joshua Boots

Submitted by:

Mathematica Policy Research, Inc.
600 Maryland Ave. S.W., Suite 550
Washington, DC 20024-2512
Telephone: (202) 484-9220
Facsimile: (202) 863-1763

Project Director:
Steven Glazerman

ACKNOWLEDGMENTS

The authors would like to thank several people who helped improve this report. From the American Board for Certification of Teacher Excellence (ABCTE), Joshua Boots and David Saba provided data, read drafts, and offered advice. Ross Rubenstein and other participants in the American Education Finance Association's annual research conference, where this report was originally presented, asked useful questions and provided useful feedback.

We thank our colleagues at Mathematica, who helped in many ways. Paul Decker, Duncan Chaplin, and Gail Baxter read the early drafts with great care. Dr. Baxter led the principal survey with assistance from Brian Roff. Elizabeth Seif provided research assistance. The report was edited by Daryl Hall and Caitlin Johnson, and Donna Dorsey provided production support.

CONTENTS

Section		Page
A	INTRODUCTION.....	1
	1. Professional Licensure and Certification	1
	2. Teacher Licensure and Certification: Traditional Approaches and the American Board Alternatives.....	1
	3. Research on Teacher Certification	2
	4. Research Questions for the ABCTE Evaluation	3
B	ABCTE PASSPORT TO TEACHING: EARLY IMPLEMENTATION FINDINGS	4
	1. The Certification Pipeline	4
	2. Passport Holders and Teachers.....	7
	3. Schools in Which Passport Holders Teach	9
	4. Principal Ratings of Passport teachers	10
C	IMPACT EVALUATION OF PASSPORT TEACHERS	13
	1. Phase I: Generic Teacher Comparisons	13
	2. Phase II: Within-School Comparisons	14
D	PREDICTIVE VALIDITY OF ABCTE EXAMS	16
	1. Novice Teachers: Praxis.....	16
	2. Veteran Teachers: National Board	18
E	DISCUSSION	19
	REFERENCES.....	21

A. INTRODUCTION

Education policymakers have long sought to establish teaching standards and measure new and continuing teachers against these standards, but existing methods for certifying teachers have been criticized for being either so onerous as to deter good candidates or so lax as to keep weak teachers in the profession. To provide another approach, the U.S. Department of Education recently funded an organization called the American Board for Certification of Teacher Excellence (ABCTE, or American Board) to develop portable teacher credentials that would identify qualified beginning teachers and recognize accomplished veteran teachers.

In 2005, ABCTE selected Mathematica Policy Research, Inc. (MPR) to conduct an independent, five-year evaluation of the certification programs. This paper describes results from the first six months and plans for the remainder of the evaluation. The first sections provide background on certification processes in use today (“traditional”) and ABCTE itself; subsequent sections discuss findings from the early use of American Board certification, as well as our design for a rigorous evaluation of the credentials, the results of which will be included in future reports. By issuing this report on the early descriptive findings and design of the evaluation, we hope to promote transparency in our approach and invite comments from the field.

1. Professional Licensure and Certification

There is longstanding debate about the benefits and costs of the regulation of a range of professions, from teaching, medicine, and law to truck driving (Cox and Foster 1990). For these and many other professional services, governments use *licensure*—restrictions on who may practice, with the goal of ensuring minimum quality and safety standards. Licensure requirements typically refer to training and examinations, sometimes a period of apprenticeship. Often, the training must be completed through an accredited institution. However, such regulations come at a cost because they create barriers to entry into these professions that may deter some who would excel in their fields. This is especially true if the requirements are imperfect predictors of quality and/or safety.

One response has been to seek non-governmental *certification*, a voluntarily sought label confirming that workers have certain skills or accomplishments. Certification has the benefit of being able to be used in compensation, hiring, promotion, or subject assignments, as well as marketing to consumers, but it is a more general term than licensure, encompassing credentials that are not mandated by law as a condition of practicing and those that are granted by non-governmental organizations. A certification process that is effective at measuring quality and has low implementation costs would be a good candidate for adoption by government regulators as a tool for licensure and by employers as a tool for hiring and promotion.

2. Teacher Licensure and Certification: Traditional Approaches and the American Board Alternatives

Both teacher licensure and certification are coming under increasing scrutiny in policy debates about teacher quality. Initial licensure for new teachers in the United States is governed by state law, which typically requires coursework in pedagogy and subject matter as well as student teaching experience and a written exam. For experienced teachers, an independent

body—the National Board for Professional Teaching Standards (NBPTS)—introduced “National Board certification” in 1987 to recognize accomplished teaching; it has since become a widely used standard. For a \$2,300 fee, the National Board analyzes detailed portfolios, including videos of classroom lessons and a standardized written exam, in order to identify and certify accomplished teachers.

Challenging both of these longstanding approaches to licensure and certification, ABCTE relies far more on written exams to determine whether prospective teachers are qualified to enter the teaching profession. The American Board is also planning to introduce a higher-level certificate to identify accomplished veteran teachers. With a lower cost and no portfolio or prior teaching requirement, these two certificates (the Passport to Teaching, or “Passport,” and a Master Teacher certificate¹) represent a new approach to certification that the American Board hopes will remove unnecessary barriers to entry while ensuring high quality (Walsh 2002).

ABCTE was founded in 2001 to help the nation’s educational entities meet the No Child Left Behind legislation’s requirement that every classroom have a “highly qualified” teacher. The organization’s mission is to increase the supply of these teachers by providing a more cost-effective and expeditious alternative to traditional, often state-specific certification. By basing its certification almost solely on demonstrated mastery of subject matter and professional teaching knowledge as measured by performance on a series of exams, ABCTE endeavors to encourage career-changers to become teachers.

In addition to high standards and cost effectiveness, American Board certification is designed to be nationally recognized—and therefore portable from state to state—as well as content area-specific. The Passport is currently recognized in five states: Florida, Idaho, New Hampshire, Pennsylvania, and Utah; ABCTE is actively working to expand this list. Initially, the Passport was offered in the areas of elementary education (grades K-6), secondary mathematics (6-12), and secondary English/language arts (6-12). More recent additions include secondary biology (6-12), secondary general science (6-12), and special education (K-12); additional subjects are in development, including chemistry and physics.

3. Research on Teacher Certification

The existing research on teacher certification is encouraging for proponents of alternatives, but the most definitive empirical evidence on whether these alternatives are effective is still being collected. The justification for creating alternate routes comes from recent empirical literature questioning the link between traditional certification and student achievement (Ehrenberg and Brewer 1994; Goldhaber and Brewer 2000) and suggesting that qualified applicants are discouraged from entering the profession (Ballou and Podgursky 1996). Policy responses to this literature have ranged widely, and have included attempts to intervene at multiple points along the route into the classroom—from recruitment and screening to training,

¹ The American Board’s program to identify, recognize and certify experienced Master teachers is still under development as of this writing (May 2006) and receives less attention in this paper. It will use the same exams as the Passport certificate, discussed below, and will require evidence of candidates’ past contributions to student achievement (“value added”).

placement, and ultimately licensure. For example, Troops to Teachers is a program that mainly provides placement services for uniformed service members seeking teaching positions. Teach for America is a program that recruits, screens, trains, and places recent college graduates from fields other than education into schools in very high-poverty areas where they commit to two years in the classroom. There is emerging research on these and related programs, comparing teachers from the different routes (Boyd et al. 2005).

Very few programs or organizations offer pure certification as ABCTE does. ABCTE relies more on self-screening and leaves preparation up to the individual. Even ABCTE, however, has found itself conducting active outreach and providing its candidates with mentors and preparation materials, although it is mainly self-guided exam preparation as opposed to traditional teacher training.

As noted above, NBPTS is another organization that offers similar certification, but for experienced teachers. There have been numerous recent studies on National Board certification (Goldhaber and Anthony 2004; Cavalluzzo 2004; Vandervoort et al. 2004). Several others are in progress. These studies typically examine the student achievement gains for certified and non-certified teachers, including failed applicants and non-applicants to the program. These studies are nonexperimental and rely on student and teacher background variables as controls to adjust for any differences in test scores between certified and non-certified teachers that are caused by factors other than certification status or teacher quality.

The experiences of school administrators and other education policymakers with National Board-certified teachers, together with recent formal evaluation research, provide a picture of the Board's effectiveness at identifying good teachers. While some have criticized NBPTS by pointing out that the certification effects are small (Cunningham and Stone 2005), early findings generally suggest that the effects are positive. We hope to leverage this information in part by examining whether National Board certification status can be predicted by performance on the American Board certification tests.

4. Research Questions for the ABCTE Evaluation

MPR's evaluation seeks to address three sets of research questions. The first set of research questions aims to help us understand the impact findings by describing how the ABCTE programs work. This report provides some early findings for the Passport to Teaching program. In future reports, we will discuss the Master Teacher certification program. The research questions include the following:

- Who chooses to apply for ABCTE certification?
- Who makes it through the pipeline to become certified?
- Where do these certificate holders end up teaching?
- How do principals view ABCTE teachers?

The second set of questions, which we will address in the coming months and years, concerns the impact of the teachers who receive the Passport to Teaching certificate. The research question is: *Are American Board-certified beginning teachers effective in the classroom?* In particular, we will examine impacts on student achievement. The first task for this analysis is to identify the appropriate benchmark for comparison. Should we compare Passport holders to beginning teachers with traditional certification or with a mix of teachers that include other alternatively certified or emergency certified teachers? When administrators prepare to hire a teacher, they might wish to know whether the average American Board-certified teacher will be as effective as the average teacher with traditional certification (i.e., the type of teacher with whom they are more likely to be familiar). We may also compare American Board-certified teachers with a mix of other teachers, including both traditional and alternatively-certified teachers.

The third and final set of research questions will address the predictive validity of American Board certification. Specifically, we pose the following research question: *Does the credential accurately identify accomplished teaching?* This is especially relevant for the Master Teacher program, but we will also explore the predictive validity of the Passport certificate. In particular, we will compare the ability of American Board certification to predict teachers' future value-added to student achievement with other widely used indicators of teacher qualifications, such as degrees, years of experience, principal ratings, and other teacher exams. In the case of the Master Teacher program, we will compare American Board certification to National Board certification.

B. ABCTE PASSPORT TO TEACHING: EARLY IMPLEMENTATION FINDINGS

At this point in the study, relatively few people have attained a Passport certificate and started teaching (about 56) and none have attained a Master Teacher certificate, as that latter program is still under development. Therefore, we have begun by gathering data on the Passport program's early implementation, since we are only now observing the first cohort of certificate holders entering the classroom.

In this section we discuss early findings from our study. These findings are based on data obtained from the ABCTE program staff, from discussions with some ABCTE teachers, and from a survey of the principals of the schools where they teach conducted in spring 2006.² We begin with a discussion of the certification pipeline, and provide early data and context for understanding who obtains ABCTE certification, and where they end up teaching.

1. The Certification Pipeline

ABCTE began certifying its first cohort of teachers in 2004, so we are only now beginning to observe them in the classroom. Although the certification was created in 2001, it took time to develop and refine a set of exams and certification processes. It is also worth noting that, as with

² As of this writing, the principal survey was still in the field, so we describe the methods but do not report findings.

many such programs, fewer people are likely to complete the process and obtain a teaching position than the number who express initial interest—many will drop off at each stage, from initial application to taking exams to completion.

ABCTE designed its Passport certification to be a quicker, more cost-effective process for highly qualified individuals than traditional certification routes. A prospective Passport holder (candidate) follows the four steps below.

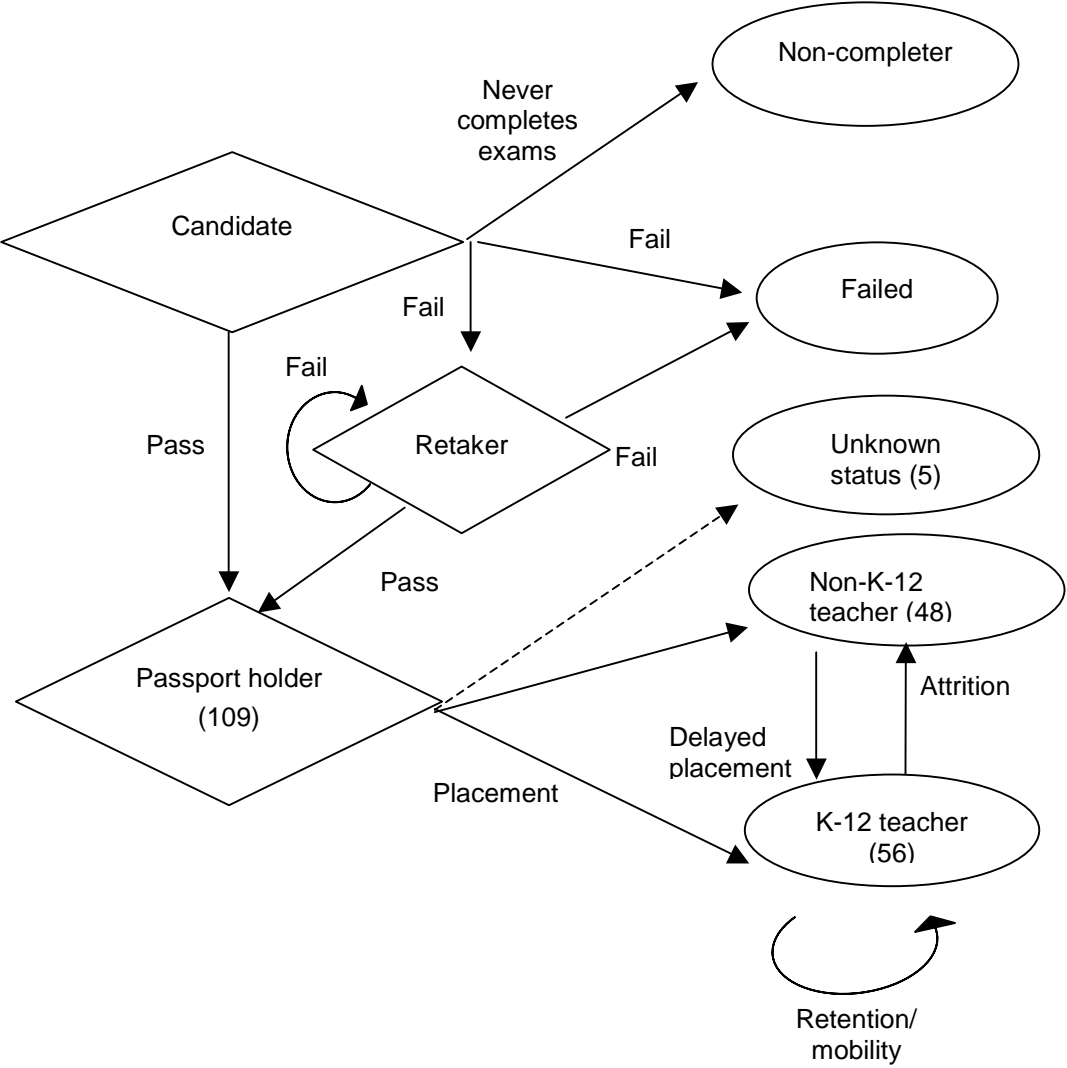
1. Enrollment: the candidate selects a certification area, requests transcripts from his or her undergraduate institution, authorizes a background check, and pays the enrollment fee of \$500.
2. Assessment: the candidate is assigned an American Board advisor who provides guidance on the process, and takes a self-assessment to identify his or her preparation needs.
3. Preparation: the candidate develops an individualized learning plan along with his or her advisor, and prepares for the certification exams using available preparation materials.
4. Certification: the candidate registers for and takes the ABCTE exams at one of the testing centers across the country. The exams include a three-hour professional teaching knowledge test (multiple choice and essay) and a four-hour subject area knowledge test. These exams need not be administered on the same day.

Candidates have up to one year to complete the entire process and may retake each exam twice. Using data provided by the American Board, we estimate the average to be eight months for the first group of candidates.

Between the beginning of the program and November 2005, 1,076 candidates paid the fees to become eligible for American Board certification. After an individual has applied to pursue ABCTE Passport certification, he or she has two possible final outcomes: meeting or failing to meet the requirements to obtain that certification (see Figure 1). Over this period, 109 people met the requirements. The remainder can be grouped into candidates who enroll but never take the exam (non-completers), those who take the exam(s) and fail (failed), and those who are still in the pipeline—some of whom may have failed the exams once or twice (retakers). According to ABCTE, roughly 70 percent of candidates who have taken the exams to date have passed.

Elementary certification is the most common. Individuals who were still enrolled in the Passport program in October 2005 represent our best prediction of the next cohort of Passport holders. Of the more than 600 such enrollees for whom we have data, 42 percent are pursuing certification in elementary education (Table 1). The rest are pursuing certification in secondary-level subjects: 25 percent in mathematics, 20 percent in English/language arts, 5 percent in general science, and 5 percent in biology. The final 2 percent are registered in special education. This distribution largely reflects the timing of rollout of each type of certification, with elementary education starting in 2003, English/language arts and math in 2004 and the other subjects in 2005.

FIGURE 1
 ABCTE CANDIDATE PIPELINE



Note: Numbers in parentheses are estimates as of 11/1/2005. Data on the numbers of candidates from the same period who never became Passport holders (including re-takers, non-completers, and failed candidates) are not yet available.

The geographic distribution of Passport candidates reflects the areas where ABCTE has gained recognition and has conducted its most intensive recruiting: Idaho, Florida, and Pennsylvania. The first group of candidates who obtained a Passport to Teaching (discussed below) came mostly from Idaho. Of the candidates who were still enrolled in October 2005, 24 percent were from Idaho, 19 percent from Florida, and 14 percent from Pennsylvania. An additional 23 percent were located overseas, which points to a large source of demand for alternative certification. The remaining 20 percent came from several other states.

2. Passport Holders and Teachers

The remainder of this section focuses on the first cohort of Passport *holders*, defined as having been certified through November 1, 2005.³ Because the Passport is geared toward career changers, these individuals look very different from other entry-level teachers in terms of demographics. Information on these individuals is based on data kept by ABCTE and augmented by interviews conducted by MPR in February 2006. Over one-third of Passport holders are male, and the mean age is 39 (see Table 1). As noted above, the population is concentrated in a few particular states—not surprisingly, those that first accepted American Board certification: over 80 percent of Passport holders reside in Idaho, Florida, or Pennsylvania.

The distribution of subjects in which Passport holders are certified is similar to the distribution of subjects pursued by candidates. Over half of the Passport holders (57 percent) obtained their certification in elementary education, followed by secondary English/language arts (25 percent), secondary mathematics (17 percent), and biology (2 percent). Passports in biology and general science were not offered before August 2005.

Once candidates have been certified, what happens to them? A slight majority of the first cohort of Passport holders (56 of 109) is currently teaching in a K-12 position in the United States (Table 2). Another 6 percent are teaching either overseas or in a postsecondary institution. An interesting finding is that some of the K-12 teachers were already employed as teachers or administrators before seeking American Board certification, although the exact number is unknown because it was not a question asked of all Passport holders. This may indicate that in its early stages, ABCTE certification is being viewed as an additional credential rather than authorization to teach.

The remaining 38 percent of Passport holders are not teaching in a full-time capacity, for a variety of reasons: 10 percent are working in an education-related field but are not in the classroom, and 6 percent are substitute teaching. The majority of the non-teachers indicated some desire to be teaching—10 percent indicated that they could not find a position; of these, over half specified that it was because the hiring authority would not accept the American Board certification (either the state, locality, or school would not accept it, or they required additional credentials).

³ Many teachers certified through this date were teaching in the 2005-2006 school year.

TABLE 1
 CHARACTERISTICS OF FIRST COHORT
 OF PASSPORT HOLDERS AND PASSPORT TEACHERS

Cohort Characteristics	Cohort 2 Candidates (%)	Cohort 1 Passport Holders	
		All Passport Holders (%)	Passport Holders Teaching K-12 (%)
Male	N/A	37.6	46.4
Age (mean years)	N/A	39	38
Location			
Idaho	24.1	55.1	60.7
Pennsylvania	13.7	21.1	19.6
Florida	19.4	5.5	10.7
Other states	19.5	15.5	9.0
Other countries	23.3	2.8	0.0
Subject Area Certified			
Elementary education	42.2	56.9	42.9
English/language arts	20.3	24.8	30.4
Math	25.0	16.5	23.2
Biology	4.7	1.8	3.6
General Science	5.3	—	—
Special Education	2.5	—	—
Total	644	109	56

N/A = not available

Overall, Passport holders who are teaching look similar to the entire population of Passport holders, with some minor differences. Table 1 shows that a slightly higher proportion of Passport holders who are currently teaching in a K-12 position (“Passport teachers”) are male, located in Idaho or Florida, and are certified in one of the secondary-level subjects (English/Language Arts, Math, or Biology).

Beyond meeting the minimum requirement for passing, performance on the ABCTE assessments does not appear to predict whether a candidate will eventually be teaching (see Table 2). The mean test score on the professional teacher knowledge test for all Passport holders was 314 out of a maximum of 500; 270 is the minimum required to pass. Of all the placement subcategories of Passport holders, those who are not currently teaching and specified that the American Board certification was not accepted for their desired position had the highest score, followed by non-teachers working in an education-related field, although both of these are small samples and may not be representative of future cohorts of Passport holders. Teachers in private schools scored slightly higher than those in public schools. Patterns were very similar for scores on the subject knowledge tests—the mean was 316 out of a max of 500; a score ranging from 252 for biology to 285 for the multiple subject exam is required to pass (this varies by subject).

TABLE 2

STATUS OF FIRST COHORT OF PASSPORT HOLDERS AS OF DECEMBER 2005

Status	Number of Teachers	%	PTK Score (mean)	Subject Score (mean)
Teaching K-12				
Catholic school	4	3.7	320	345
Other nonpublic school	8	7.3	315	313
Other (presumably public)	44	40.4	310	311
All	56	51.4	311	313
Teaching non-K-12				
Overseas	3	2.8	285	332
Postsecondary	4	3.7	317	313
All	7	6.4	301	323
Not Teaching				
Still in education-related field	11	10.1	325	342
Substitute teaching	6	5.5	312	312
Could not find job	5	4.6	300	305
ABCTE not accepted	6	5.5	334	329
Deceased	1	0.9	—	—
Other/reason not reported	12	11.0	318	313
All	41	37.6	319	322
Unknown Status	5	4.6	—	—
All Passport Holders	109	100.0	314	316

PTK = Professional teaching knowledge

3. Schools in Which Passport Holders Teach

In addition to understanding the characteristics of the first cohort of Passport holders and teachers, it is useful to understand their teaching assignments, as well as the characteristics of the schools in which they teach. Using the National Center for Education Statistics' (NCES) Common Core of Data (CCD) from 2003-2004, we were able to observe characteristics of these schools, although it is important to keep in mind that the relatively small sample size of 45 precludes making definitive statements about patterns.

Passport holders' teaching assignments are similar to what we would expect, based on the characteristics of Passport holders in general (Table 3). One exception to this is that we observe fewer teachers in elementary schools, relative to the number of Passport holders certified in elementary education. In fact, almost 30% of Passport holders are not teaching in their certified subject area. It is unclear whether this is due to the lack of open positions for the Passport holders in their certified area, or if these teachers pursued certification in the subject matter that most closely matched their desired field of instruction. In almost 15 percent of the schools, Passport holders are teaching in a subject area that is not certified by ABCTE (examples include drama, special education, and foreign language).

TABLE 3
CHARACTERISTICS OF PASSPORT HOLDERS' TEACHING ASSIGNMENTS

School Characteristics	ABCTE Schools (%)
Location	
Idaho	60.0
Pennsylvania	22.2
Florida	11.1
Other states	6.7
Grade Configuration	
Elementary	17.8
Middle	8.9
High	35.6
Other	37.7
Subject Area Taught by ABCTE Passport Holder	
Elementary education	26.7
English/language arts	20.0
Math	28.9
Science	11.1
Other	13.3
Teaching in certified subject	73.2
Total	56

Passport holders teach in schools that are similar to other schools in the locations where they teach (see Table 4). The schools serve predominantly white students: 80 percent, which is much higher than 60 percent nationally but lower than 86 percent in Idaho (where the majority of Passport holders are teaching). Similarly, the proportions of schools that are in urban or rural areas differ from the national average but lie between the corresponding proportions in Idaho and Pennsylvania. Charter schools also seem to be heavily represented in the population of ABCTE schools, as do schools with a schoolwide Title I program. Combined with the large proportion of rural schools, these patterns may indicate that schools in higher-need locations are more likely to have openings for American Board-certified teachers.

4. Principal Ratings of Passport teachers

In order to provide some early evidence on the performance of Passport holders before the test score data become available (see section C below), we administered a survey of the supervisors, usually school principals, of the first cohort of Passport holders who have teaching positions in the 2005-2006 school year. The survey is still in the field as of this draft, so we describe the methods here. The results will be available in June 2006.

TABLE 4
CHARACTERISTICS OF SCHOOLS IN WHICH PASSPORT HOLDERS TEACH,
REPRESENTED STATES, AND NATIONWIDE

School Characteristics	Percentage of Schools			
	ABCTE	Idaho	Pennsylvania	U.S.
Population White	79.2	85.9	77.1	59.5
Population Eligible for Free/Reduced-Price Lunch	31.3 ^a	37.0	28.1	35.3
Title I School (schoolwide)	35.6	13.2	15.8	25.0
Type of School:				
Traditional public	77.8	86.7	60.9	75.6
Charter school	2.2	0.1	1.5	1.9
Private	15.6	12.0	37.7	22.4
Locale ^b				
Urban	22.2	15.3	19.0	24.6
Suburban	28.9	6.8	40.1	32.4
Rural	48.9	77.9	40.9	43.0
Number of Schools	45^c	782	5,215	121,335

Note: Data on state and national rates from Common Core of Data (CCD).

^aThe population eligible for free/reduced-price lunch is calculated as the average of the school averages for the ABCTE schools. Because the data are not readily available at the state or national level, those are student-level averages.

^bLocale is based on 2003-04 CCD district codes as follows: “urban” districts include large city or mid-size city; “suburban” districts include urban fringe of a large or mid-size city; and “rural” districts include large town, small town or rural.

^cThere are fewer than 56 (the number of passport holders teaching K-12) because 5 schools were not located in the CCD, 5 schools did not have available data in the CCD, and 1 school was not identified by the teacher.

Principal ratings are subjective, yet they represent one of the closest perspectives on a teacher’s performance by someone whose job is to assess and judge the quality of the teacher’s work. Recent research has shown principal ratings to be fairly good predictors of student achievement, at least at the upper and lower tails of performance (Jacob and Lefgren 2005).⁴ In addition, because our ratings instrument is not specific to tested subjects and grades, the method allows us to judge teachers across the spectrum of grades and subjects.

⁴ We will conduct a similar analysis in the future as part of the current evaluation, with the advantage of including detailed teacher exam scores. The sample size will be small for the first cohort of teachers, but will grow as we add succeeding cohorts.

The respondents were asked to rate teachers on a five-point scale on the following dimensions of teacher performance:

- Understanding student needs
- Planning lessons
- Leading instructional activities
- Adapting instruction during a lesson
- Managing the classroom
- Encouraging desired student behavior
- Engaging students in learning
- Communicating content knowledge
- Collaborating with other teachers
- Responding to feedback from teachers and administrators
- Relating to parents.

We asked the supervisor to rate the teacher on each of these dimensions of performance as well as overall performance compared to two benchmarks: all teachers they have observed in their careers and other first-year teachers they have observed. Responses are rated on a five-point scale. We also ask about the Passport holders' potential for developing into an effective teacher in the future.

We administered this survey to supervisors of the universe of Passport holders who teach students in K-12 schools in the United States, regardless of public/private sector location. In 2005-2006, there are 56 such teachers.⁵ In future years, we will administer the survey to the much larger cohorts that are expected to enter the classroom starting in the fall of 2006, focusing on a stratified random sample of teachers if appropriate.

In addition to tabulating the overall ratings of teachers, we intend to examine the correlation between teacher ratings and teacher scores on the American Board exams as well as regression-adjusted student test score gains, which approximate the value added by the teacher to achievement (see below). We will test the hypothesis that higher scores predict stronger performance as judged by principals.

⁵ In some cases a principal may supervise multiple teachers. Findings from the first cohort of Passport holders suggest that this is rare.

C. IMPACT EVALUATION OF PASSPORT TEACHERS

The next major research question the evaluation will address is: Are American Board-certified beginning teachers effective in the classroom? This section summarizes our approach to testing the hypothesis that Passport holders are at least as effective in the classroom as their traditionally licensed counterparts. Findings from this component of the evaluation should become available beginning in late summer 2006, following the release and analysis of spring 2006 test scores for the initial cohort of Passport teachers.

The research design proceeds in two phases. First, to reflect the need for timely information, Phase I will compare outcomes for teachers from different schools, including large pools of comparison teachers. To provide more rigorous evidence, Phase II will compare teachers within the same schools and grades using random assignment of students to classrooms. This type of evidence will require participating schools' active cooperation and take multiple years in order to accumulate a large enough sample to draw firm conclusions. The main outcome of interest in both phases is the teachers' productivity, or "value added" to student achievement. We will estimate this outcome by using student test score data and, where available, student background information, which includes prior test scores. Phase II will seek to confirm the findings in Phase I. If the two sets of estimates differ, we will reconcile the findings by exploring whether American Board certified teachers landed more difficult teaching assignments, whether impacts vary considerably by classroom and school, or some other explanation.

1. Phase I: Generic Teacher Comparisons

The goal of the first phase of the study is to provide early evidence on the effectiveness of Passport teachers in the classroom. For teachers in tested grades and subjects, we will collect student test scores and estimate the growth in achievement attributable to whether the teacher was certified by ABCTE or not. The unique contribution of the teachers to student performance is their value added.

The statistical model for estimating value added is straightforward. Using a sample of students in the classrooms of study teachers and those from a group of comparison teachers drawn from similar schools, we will estimate the following achievement growth model:

$$Y_i^t = aY_i^{t-1} + b'X_i + \sum_j cSchl_{ij} + dABC_i + u_i$$

where:

Y = test score for student i in year t ⁶

X = vector of student characteristics (race/ethnicity, sex, free lunch eligibility, language, and disability)

$Schl$ = indicator set to 1 if student i attends school j , else zero

⁶ Noting that pre-test (Y^{t-1}) is measured with error, we correct for errors in variables using an estimate of the reliability of the test, which we will obtain from the publisher, NWEA.

ABC = indicator set to 1 if student i is taught by Passport holder, else zero
 a, b, c, d = parameters to be estimated
 u = error term, assumed to be independent and identically distributed across individuals

The test scores are the Measure of Academic Progress (MAP) developed by the Northwest Evaluation Association (NWEA). MAP math, reading, language, and science scores for grades 2 through 12 will be available for nearly all schools where Passport holders are teaching in 2005-2006. In addition, the scores are linked by unique student ID over time so that we can use more than one time point for a pre-test measure if needed. The goal of the analysis is to test whether coefficient (d) on the American Board indicator is less than zero. If $d=0$, then Passport holders are as good as (and not worse than) regular certified teachers.

Will these estimates truly reflect the “value added” by the Passport holders—in other words, will d be a reliable, unbiased impact estimate? The two concerns are sample size and omitted variables. Sample size is a limitation because, although we can draw a very large comparison group of teachers, there is a fixed number of American Board-certified teachers in the classroom as of 2006, and an even smaller group who are teaching in tested grades and subjects. (We estimate that there are about 18 such teachers in the first cohort.) We plan to address this concern by accumulating more sample in future years. In 2007, we expect there will be a much larger sample of available certified teachers, primarily in Florida. The Florida Department of Education maintains an education data warehouse of comprehensive linked teacher and student data, which we expect to tap for this effort.

Omitted variables should always be a concern because there is no guarantee that the students who are in classrooms taught by Passport teachers do not differ from other students in some way that cannot be captured by the model’s control variables. For example, American Board-certified teachers may choose to work in schools that are harder to staff and have more challenging working conditions, such as unmotivated or unruly students. We plan to address these concerns in Phase II, discussed below.

Findings from the first round of value added analysis should be available in late summer of 2006, after the tests are administered in the spring of that year.

2. Phase II: Within-School Comparisons

To strengthen the estimates of the Passport impact, we plan to conduct comparisons within schools using random assignment of students to teachers where feasible. Our Phase I study of the impact of Passport teachers on student outcomes has the same limitation of the studies of National Board certification cited earlier (Goldhaber and Anthony 2004; Cavalluzzo 2004; Vandevort et al. 2004). That is, it relies on observable control variables in a linear regression to account for all the factors that might also explain differences in student achievement beyond certification status. The findings rest on an untestable assumption that we have accounted for all the factors that matter, including unobservable determinants of student achievement. The random assignment strategy will ensure that we have a set of teacher pairs (or teacher groups) with nearly identical sets of students on average. We will obtain lists of students entering the grade or subject where there is more than one eligible teacher, including a Passport holder and a traditionally certified teacher, and randomly assign students to each teacher's roster.

Random assignment within schools on a large scale presents challenges, but it has been done successfully (Glazerman et al. 2006). It requires the cooperation of schools and must be arranged well ahead of time, before students are assigned to classrooms—preferably in the spring before the school year under study. However, given our interest in including specific types of teachers, we must also know when recruit schools for the study where those teachers will be working and what their teaching assignments will be. Thus, we intend to conduct the Phase II analysis using Passport holders who are returning for their second year of teaching, which would normally be in the same assignment as their first year.

Within-school comparisons are especially strong because they hold constant all the variables that are associated with the school, such as the quality of the leadership, the neighborhood, and the general student characteristics. Random assignment of students assures that there will be no systematic relationship between student characteristics and the “treatment” status (Passport holder or traditionally certified teacher). Prior arrangements must be made to ensure that principals restrict student transfers between classrooms so that the “treatment” is not contaminated. In many cases, roster checks are useful to verify that this policy was enforced and document any deviations so they can be accounted for in the analysis. Where random assignment is not feasible, we will document the process by which students were assigned to classrooms and the reasons that the researcher was unable to control the student assignments.

One limitation of this approach is that it requires us to focus on a narrow subset of Passport holders and therefore limits our ability to estimate impacts for all American Board-certified teachers. Passport holders can only be eligible for Phase II if they are returning to teach in a known school and grade—in other words, they are not first-year teachers—and they teach in tested grades and subjects in schools where there is at least one other person teaching the same subject to a similar group of students. These criteria eliminate Passport holders who leave after one year, teach non-core subjects, or teach in early elementary grades, small schools, or schools that have extensive team teaching or looping, a practice whereby students stay together and with the same teacher from year to year.

Another limitation is that certification status (Passport or traditional certification) may be confounded with experience level. The traditionally certified teachers that we find in the same schools as Passport holders will include some mix of novices and veterans, whereas the Passport holders in this phase of the study will mostly be second-year teachers. Our preference might be to compare Passport holders to other teachers of the same experience level—in other words, second-year teachers—but it may be difficult to find enough schools where such comparisons are possible, and even more difficult to screen schools ahead of time to identify which ones would provide such novice-to-novice comparisons. Nevertheless, we will examine the size of the future cohorts and determine the extent to which we can screen and over-sample schools that are more likely to have newly certified non-Passport teachers alongside Passport teachers. To the extent that our comparisons confound teacher experience with certification status, we can bound the size of the Passport effects by assuming that performance of experienced traditional teachers is no worse than that of beginning traditional teachers. Thus, if we empirically reject the hypothesis that Passport teachers are no worse than the average traditionally certified teacher, we can conclude that they are also as good as or better than the beginning traditionally certified teacher.

Despite these limitations, within-school comparisons of teacher effects will provide a strong test of the program's ability to attract and certify teachers who are as good as or better than those already deemed by the state to be highly qualified by traditional criteria.

D. PREDICTIVE VALIDITY OF ABCTE EXAMS

The previous section discussed our approach to estimating the impact of teachers certified by ABCTE's initial certification program, the Passport to Teaching. Because American Board certification is based almost exclusively on the results of two teacher exams, studying the examinations themselves offers another test of their usefulness. In this section, we discuss our approach to estimating the ability of the exams to predict outcomes on other widely used teacher certification exams such as the Praxis series published by the Educational Testing Service (ETS).⁷

Broadly speaking, we want to know whether the ABCTE exam to certify beginning or master teachers is a good predictor of performance in the classroom. An indirect way to address this question is to determine whether passing the American Board certification exam is a good signal of teacher quality or qualification. We can test that hypothesis by comparing it to other signals of teaching quality that have already been adopted by policymakers.

The Praxis exams, particularly the Praxis II, are used in some way by most states in their teacher certification process. However, they are typically just one indicator and passing the exams is a necessary but not sufficient criterion for licensure. Therefore, it represents a useful benchmark against which many state policymakers would like to judge the American Board exam.

1. Novice Teachers: Praxis

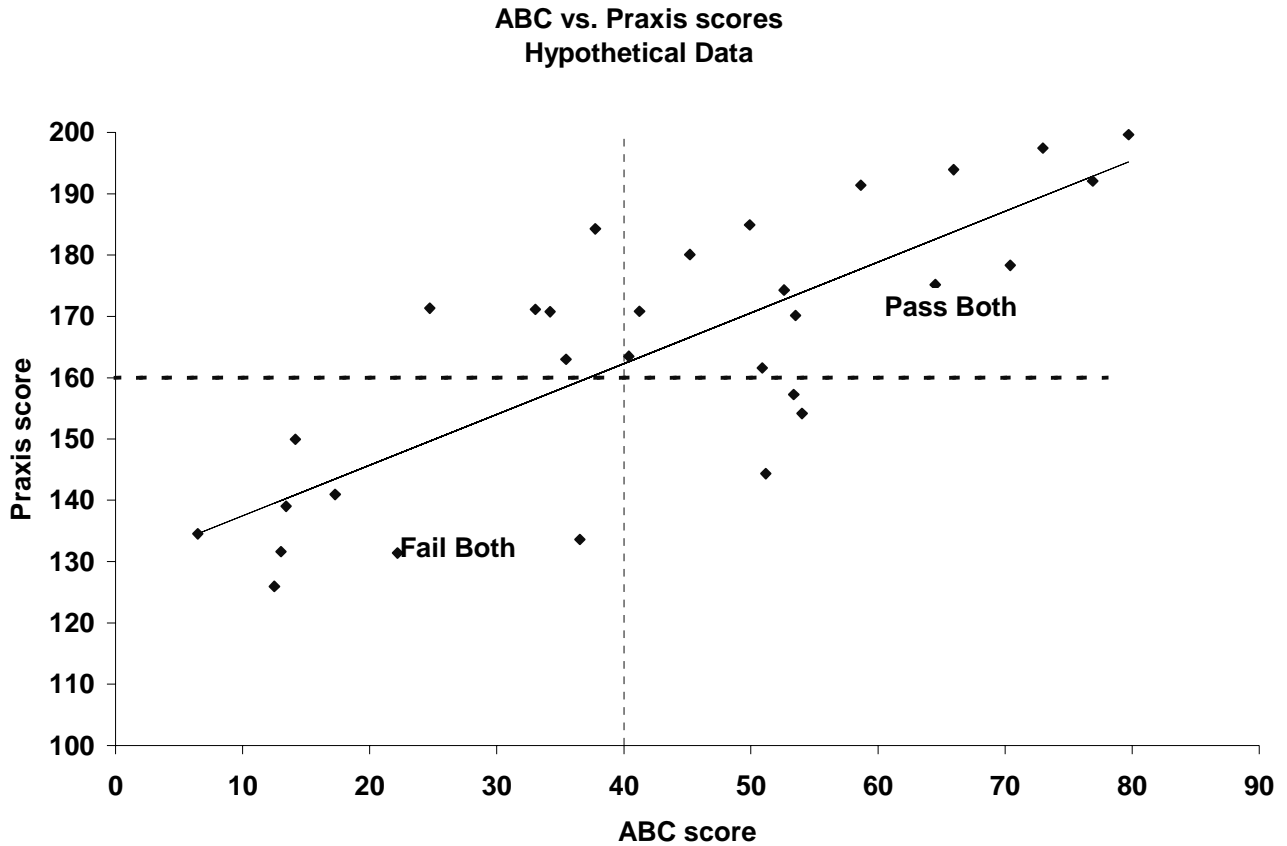
The research question we aim to address is: How well does performance on the American Board exams predict performance on the Praxis II exams? The obvious way to test this hypothesis is to select a random sample of the population of American Board candidates, administer the Praxis II under conditions that are as realistic as possible. The data that we would be trying to model can be illustrated very simply with a plot. Figure 2 shows a hypothetical example with made-up data to illustrate the relationship we are trying to estimate. A strong positive correlation would suggest that the exams measure the same underlying construct. Once we collect the exam data and background data on teachers, we will estimate a regression model such as the following:

$$PRAXIS = gABC + h'W + v$$

⁷ The Praxis series of teacher exams includes the following: Praxis I targets prospective teachers earlier in their careers, before they enter teacher education programs; Praxis II exams, which cover subject matter knowledge as well as teaching foundations and principles, are mainly used in evaluating and certifying prospective or beginning teachers; Praxis III is primarily a classroom observation measure used to assess those already teaching.

FIGURE 2

ILLUSTRATION OF HYPOTHETICAL FINDINGS FROM PRAXIS VALIDATION STUDY



where:

PRAXIS = score on the Praxis II Principles of Learning and Teaching (PLT) exam

ABC = score on the American Board's Professional Teacher Knowledge (PTK) exam

W = vector of teacher characteristics that are observable to an administrator who would be hiring the teacher

g = estimated change in Praxis II PLT score per unit difference in American Board PTK score

h = vector of coefficients relating teacher characteristics to Praxis II PLT scores

v = unobserved random error component, assumed to be uncorrelated with *ABC* and *W*

We will also estimate more general versions of this model, allowing for nonlinear relationships between *ABC* and *PRAXIS*, and interactions between *ABC* and *W*. In other words, we hypothesize that very high or low *ABC* scores could have more or less predictive power than scores near the cutoff and that *ABC* scores may be more predictive of Praxis performance for some types of teachers than others. For example, age and undergraduate training will be included in *W*. We intend to conduct a teacher background survey and request consent to obtain college entrance exam score (SAT or ACT) in order to further enrich the specification of *W*.

The model above uses continuous exam scores, but we can more closely simulate the policy decision by estimating a more simplistic model that predicts a binary outcome of whether the teacher would pass the Praxis exam. In that case, we would estimate the same model as a logit or probit, with the left-hand side expressed as an indicator equal to one if the Praxis score is greater than some cut-off. On the right-hand side, we can enter ABC scores as a binary indicator as well as continuous, or as an interaction of the binary and continuous.

Finding enough Passport holders to take the Praxis exams may be challenging if the Praxis is not a requirement for licensure. A more practical design might identify a sample of Praxis II takers who would be willing to sit for the American Board exams. Large groups of Praxis II takers can be found at major schools of education. This has the advantage that recruiting would be easier but the drawback that the population might not be applicable to American Board candidates. The latter could be addressed by setting eligibility criteria to select volunteers who look more like American Board candidates based on characteristics such as age, gender, and work history.

2. Veteran Teachers: National Board

For veteran teachers, American Board certification would play a different role, one more akin to true certification rather than licensure, because it is not required to remain in teaching. Again, we can test the importance of that signal by comparing it to other, more widely accepted measures such as the NBPTS Master Teacher designation, which can be thought of as the standard. National Board certification is a time-intensive and more expensive process than American Board certification—thus, if passing the latter is highly predictive of passing former, the American Board process could be considered a more efficient alternative.

To assess the ability of American Board exams to identify accomplished teachers, we will focus on one or two states where data—ideally from multiple years—are available to construct high-quality indicators of teacher performance. These indicators will be based on student achievement data linked over time for multiple years of a given teachers' students. Such data are available in two states in particular: North Carolina and Florida, which also happen to have high concentrations of National Board-certified teachers.

We plan to recruit volunteer teachers with at least three years of experience in a specified grade or grades to take the American Board exams and complete a background survey, providing us with consent to obtain their college entrance exam scores. We plan to specify a grade and subject area—for example, eighth grade mathematics—in order to maintain a clean comparison that does not confound these criteria with the variables whose relationship we are trying to estimate: American Board scores, National Board scores, and test score-based estimates of value added. The disadvantage of focusing on a single grade and subject is that the findings would not necessarily generalize to all teachers. Therefore, it is important to focus on a grade/subject combination that is of policy interest in its own right. Middle school mathematics is a good example of a teacher assignment that is seen as high-demand.

Findings from this validity exercise using veteran teachers will be very useful. They will allow the American Board to consider the rigor of the exams and the relative weight to place on outside factors such as evidence of value added from prior years' test scores. Just as important,

the findings will allow policymakers to decide whether to use the credential as a consideration in hiring, promotion, or bonus awards.

E. DISCUSSION

This report uses very early data on the participants, the pipeline, and the schools where Passport holders teach to suggest that the program is enrolling and screening increasing numbers of candidates. The average age of nearly 40 suggests that the candidates tend to be career changers with more work experience than other entry-level teachers. The candidates are overwhelmingly located in states that recognize American Board certification and where ABCTE has concentrated its recruiting efforts, although many currently in the pipeline come from outside the U.S.

Most Passport holders were able to find teaching positions shortly after receipt of that certification; however, several of them did not. Those who did not can be grouped into those who chose not to teach (and pursued certification for other reasons or for future use) and those who were unable to find positions, even in states that recognize Passport certification. Further, several Passport holders teach at private or charter schools, which generally do not impose the same credentialing requirements as traditional public schools. The characteristics of the schools in which Passport holders teach appear to be driven by the characteristics of the states in which they reside, although there are some minor differences possibly due to chance.

Some of the early findings raise questions that we will address through additional analysis of program implementation. For example, a consequence of a low-cost credential is that some people may obtain it without intending to teach. We found some of the highest-performing exam takers going into education-related fields, but not the classroom. One possibility, therefore, is that the credential is used to bolster one's credibility in the education industry without necessarily being a gateway into teaching. Another possibility is that the first cohort of Passport holders is idiosyncratic, with a higher proportion of early adopters seeking to "try out" the new certification program. We have not yet compared the rate at which Passport holders enter teaching with the comparable rate for traditional teacher candidates, so we cannot conclude that the Passport is atypical.

Because we are still in the first year of a five-year study, much of this report discusses future research plans. The release of interim reports like this one is important for transparency of the evaluation's methods. It is also important to provide timely research evidence, even if it is based on preliminary findings, because ABCTE and its evaluator face a "chicken-and-egg" problem: the program needs early evidence for the program to be widely accepted, yet the program needs acceptance in order to generate a study population large enough for a rigorous evaluation.

Our research strategy in the face of this challenge is to build up an evidence base incrementally, increasing the sample size and level of methodological rigor over time. We first gathered descriptive data and are now soliciting subjective ratings of teacher performance, such as those supplied by school principals. Next we will attempt to estimate the impacts of American Board-certified teachers on student achievement by collecting test score data from students taught by American Board-certified teachers and comparing them to scores for students taught by other teachers. We begin in Phase I, where we compare scores across classrooms in different schools. Then we will proceed to Phase II, where we seek schools' permission to

randomly assign students to classrooms within a school. This method represents a controlled experiment, the gold standard in evaluation research, in which we create identical classrooms of students and can attribute differences in achievement growth to the performance of the teachers. If appropriate, we will next design a strategy for a Phase III, which will take the Phase II research design to a larger scale.

Early findings may motivate ABCTE to adjust both the certification process and the program's recruiting and outreach strategies. In addition, the pool of applicants could change over time. Together, these make extrapolation from early results to "steady state" operation problematic, but they justify the steady, cautious buildup of an evidence base from formative to summative evaluation and from small-scale to large-scale investigation.

REFERENCES

- Ballou, Dale, and Michael Podgursky. *Teacher Pay and Teacher Quality*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1996.
- Boyd, Donald, Pamela Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff. "How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement." Albany, NY: State University of New York, November 2005.
- Cavalluzzo, Linda. "Is National Board Certification an Effective Signal of Teacher Quality?" Alexandria, VA: The CNA Corporation, November 2004.
- Cox, Carolyn, and Susan Foster. "The Costs and Benefits of Occupational Regulation." Washington, DC: U.S. Federal Trade Commission, October 1990.
- Cunningham, G.K., and J.E. Stone. "Value-Added Assessment of Teacher Quality As an Alternative to the National Board for Professional Teaching Standards: What Recent Studies Say." In *Value Added Models in Education: Theory and Applications*, edited by Robert Lissitz. Maple Grove, MN: JAM Press, 2005.
- Ehrenberg, Ronald, and Dominic J. Brewer. "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond." *Economics of Education Review*, vol. 14, 1994, pp. 1-13.
- Goldhaber, Dan, and Emily Anthony. "Can Teacher Quality be Effectively Assessed?" Washington, DC: Urban Institute, April 2004.
- Goldhaber, Dan, and Dominic J. Brewer. "Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement." *Education Evaluation and Policy Analysis*, vol. 22, 2000, pp. 129-145.
- Glazerman, Steven, Daniel Mayer, and Paul Decker. "Impacts of Teach for America on Student Achievement and Other Outcomes." *Journal of Policy Analysis and Management*, Winter 2006.
- Jacob, Brian, and Lars Lefgren. "Principals As Agents: Subjective Performance Measurement in Education." NBER Working Paper No. 11463. New York: National Bureau of Economics Research, June 2005.
- Vandevoort, L.G., A. Amrein-Beardsley, and D.C. Berliner. "National Board Certified Teachers and Their Students' Achievement." *Education Policy Analysis Archives*, vol 12., no. 46, September 8, 2004. Available at [<http://epaa.asu.edu/epaa/v12n46/>].
- Walsh, Kate. "Teacher Certification Reconsidered: Stumbling for Quality." Baltimore, MD: Abell Foundation, June 2002.