DOCUMENT RESUME

ED 477 460 TM 034 990

AUTHOR	Qu, Yanxuan; Becker, Betsy Jane
TITLE	Does Traditional Teacher Certification Imply Quality? A Meta-Analysis.
PUB DATE	2003-04-00
NOTE	48p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25, 2003).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	*Alternative Teacher Certification; *Effect Size; Elementary Secondary Education; *Meta Analysis; *Preservice Teacher Education

ABSTRACT

This meta analysis examined the results of 24 studies in which traditionally certified teachers were compared with teachers with a variety of other kinds of certificates. The 24 studies allowed for the computation of 192 effect sizes, and 7 kinds of comparison were performed. Findings suggest that traditional teacher training is at least as effective as alternate route training and more effective than minimal (emergency) certification. However, it is clear that some alternative teacher training programs are equally effective in providing quality teachers, and one important predictor of differences in program effectiveness was the location at which teachers were studied (and often trained). The role of experience was highlighted in the comparisons of in-field and out-of-field teacher. In this situation, differences were not apparent for new teachers, but findings favored experienced in-field teachers. An additional finding was that the studies of these alternate routes to teacher certification vary greatly and are not always well reported. Multiple confounded study characteristics appear to relate to the magnitudes of differences that were found, but much addition that might have been of use to the analyses were not reported. (Contains 9 tables and 45 references.) (SLD)

Does Traditional Teacher Certification Imply Quality? A Meta-Analysis

Yanxuan Qu
Betsy Jane Becker

Michigan State University

Presented at the annual meetings of the American Educational Research Association

Chicago, Illinois
April 23, 2003

The issue of how to improve teaching quality has led to a strenuous debate centering on the types of qualifications we should require of new teachers. Teacher-certification requirements have been varied to increase the numbers entering the teaching profession, as well as to keep the level of teacher qualifications reasonably high in the face of teacher shortages. The current secretary of the U.S. Department of Education, Grover Whitehurst, has endorsed alternate routes to certification, and Title II of the No Child Left Behind Act of 2001 provides for two specific alternate routes (Transitions to Teaching and Troops to Teachers) that may help staff high-needs schools (http://www.ed.gov/legislation/ESEA02/pg28.html).

The specific requirements for becoming a regular or traditionally certified teacher have differed over time and, at any particular time, across locations. The question, of course, is whether teachers who have not earned traditional teaching certificates perform as well as traditionally certified teachers. To date no effort has been made to systematically synthesize the literature on alternate routes to certification. In this paper we examine the results of 24 studies in which traditionally certified teachers are compared to teachers with a variety of other kinds of certificates. We use the methods of

meta-analysis to assess the magnitudes of differences between certification groups, and to clarify the factors that lead to variation in those differences.

We begin with a discussion of the types of certification that have been examined in the U.S. and in our studies. We then describe the studies we have gathered, as well as a small set not included in our meta-analysis. We conclude with analyses and a discussion of their implications for alternate routes to certification.

Studying Teacher Certification

Since 1960 many studies have described, and compared the effectiveness of, teachers with different kinds of teaching certificates. Teachers' classroom performance and their students' achievements are often used as indicators of teaching quality. Some researchers have concluded on the basis of classroom observations or student achievement that traditionally certified teachers are more effective than emergency (or provisionally) certified teachers (Beery, 1960, 1962; Laczko & Berliner, 2000, 2001; Laczko-Kerr, 2002) and alternatively certified teachers (Laczko-Kerr, 2002).

Not all results support the superiority of the traditional route. Early on, Shim (1963) found that students taught by a series of not yet certified teachers may actually score higher than those taught by a series of certified teachers. Hawk and Schmidt (1989) compared traditionally certified teachers and alternatively certified teachers in terms of their classroom performance; their results showed traditionally certified teachers were superior on some outcomes, but alternatively certified teachers were superior on others. Dewalt and Ball (1987) also found mixed results, favoring traditionally certified teachers and emergency certified teachers on different outcomes. Miller et al. (1998) compared traditionally certified teachers with alternatively certified teachers in terms of both

teaching performance and student achievement. They concluded that teachers with traditional certificates and alternative certificates are equally effective in terms of both teaching performance and student achievement. These diverse results have encouraged us to synthesize studies of certification conducted in different grades and subject-matter areas, and using different designs, to examine the effectiveness of traditionally certified teachers and teachers with other kinds of certificates.

Types of Teacher Certificates

There are three main types of teacher certificates-the traditional teacher certificate, alternative teacher certificate, and emergency teacher certificate. Within each are variations in program activities, program length, and duration of the certification. Some authors also refer to provisional or temporary certification, which typically means a teacher has satisfied the requirements of a standard certificate but either has little or no teaching experience (or no recent experience), or has taught in a different locale. The National Association of State Directors of Teacher Education and Certification (NASDTEC) publishes an annual compendium of teacher certification requirements (e.g., NASDTEC, 2002).

Traditional teacher certificates set the greatest requirements for teachers. Teachers typically earn a bachelor's degree in education, and must have finished student teaching under the direction of a supervisor or master/mentor teacher (Brown, 1987; Cornett, 1984; Laczko-Kerr, 2002; Sandlin, Young & Karge, 1993). Alternate routes to certification often ask that participants have at minimum a bachelor's degree, but the degree need not be in education. The emergency teacher certificate has the least requirements.

Table 1 shows similarities and differences among beginning teachers with different types of teacher certificates. These features were drawn from the set of studies we analyze below. The table reveals some clear differences in requirements of these certificates. Teachers with traditional standard (or full) teacher certificates and traditional provisional teacher certificates appear to differ only in their levels of teaching experience. Similarly, out-of-field teachers meet many of the requirements set for traditional certification (indeed they may have some kind of full certification), but they teach a subject which is not covered by their certification. For instance a teacher fully certified to teach language arts could be classified as "out-of-field" if he or she were assigned to teach math or science.

Table 1. *Requirements for Beginning Teachers of Different Types of Certificates*

Type of Certificate			*Traditional*	*Out-of-field*	
	Alternative		*Emergency*		
	Standard	*Provisional*			
Bachelor's degree	Education	Education	Education	Education or subject-matter major	Subject-matter major	May not yet have degree
Education courses (while working)	Yes	Yes	Yes	Yes	None or some
Student teaching	Yes	Yes	Yes	Yes (while working)	No
Teaching experience	Yes	No	Yes or no	Yes or no	No
Program length	4 or 5 years	4 or 5 years	4 or 5 years	1 year or less	---

Alternative teacher certificates usually are issued to on-the-job teachers after they have finished an alternative teacher-training program. Such programs often involve on-the-job training, in that participants are given full-time teaching jobs where they are observed on an ongoing basis by mentor teachers. Their teaching internship is typically more intense than the student teaching completed by traditionally certified teachers.

- **Input Drive** (the teacher searches for new ideas and experiences to share with students)
- **Activation** (the teacher motivates students to think, respond, and feel in order to learn)
- **Innovation** (the teacher is determined to implement creative new ideas and techniques)
- **Gestalt** (the teacher tends toward perfectionism, but works from individual to structure)
- **Objectivity** (the teacher responds to the total situation rather than with impulsive reactions)
- **Focus** (the teacher has models and goals and selects activities in terms of these goals)

Recently, Gallup has expanded the range of interview services provided. In early 2002 Gallup unveiled the a Web-based "talent assessment system" that asks applicants to respond to a series of statements using a 5-p range of multiple-choice questions that reveal their attitudes, beliefs, and behaviors, and to a number of ope TeacherInsight takes approximately 40 minutes to complete, and the program nearly immediately generates that provides the applicant's percentile ranking of his/her predicted potential for teaching success based on responses fit with Gallup's themes. The Gallup website markets TeacherInsight as their "next generation" p "hire the best teachers...fast" (http://education.gallup.com). Potential clients are informed the system can h pool of candidates like their best teachers, using a centralized approach that requires less staff time. Further even shorter version for schools with a high volume of teacher candidates. The Automated Teacher Screen through an automated telephone interview in which they are asked to respond to a series of statements usin Gallup provides feedback to school personnel via their website, sorting the candidates into priority levels fo interviews, either the longer TeacherInsight or the actual face-to-face TPI. (http://education.gallup.com)

Gallup's Urban TPI. Perhaps in response to Haberman's Urban Teacher Selection Interview, Gallup begar Teacher Perceiver Interview in the late 1990s. The Urban TPI claims to identify the "best" urban teachers r series of consistently recurring patterns of thought, feeling, and behavior. Whereas the regular TPI is organ themes, the Urban TPI is based on eleven themes that define who successful urban teachers are and how th

- **Commitment** (the teacher consciously decides to contribute to people through education and works primarily need, in spite of obstacles)
- **Dedication** (the teacher finds satisfaction in student development and emotionally becomes a part of their lives
- **Individualized Perception** (the teacher considers the interests and needs of each student)
- **Caring** (the teacher shows warmth to students and gives priority to developing relationships)
- **Involver** (the teacher wants to be partners with their students, with parents, and with other teachers and to give education)
- **Empathy** (the teacher deals with the individual student's feelings and thoughts)
- **Positivity** (the teacher has hopeful attitudes toward students)
- **Initiator** (the teacher is an advocate for students and will speak up to make a difference)
- **Stimulator** (the teacher is personally dramatic and receptive to the ideas of students)
- **Input** (the teacher searchers for new ideas and experiences to share with students)
- **Concept** (the teacher is guided by positive learning concepts for what is best for students)

Comparing the Interviews. The two major structured interviews for teachers share some notable similaritie that both the TPI and the Haberman interviews are structured around comparable thematic frameworks. She Mission and Investment themes (which she classifies as "intrapersonal") pair up with what she identifies as Professional" theme from Haberman's original interview; Focus matches up with what she identifies as the Professional," "Theory & Practice," and "Burnout" themes from Haberman. The TPI's "interpersonal" then Drive, and Listening correspond with what she identifies as the "Approach to At-Risk" and "Professional V from Haberman; she pairs Objectivity with what she identifies in the Haberman interview as "Promoting L "Persistence." Ryan sees less clear-cut correspondence between TPI's "extrapersonal" themes of Individual Drive, Activation, Innovation, and Gestalt—each touches upon at least two of Haberman's original themes are touched upon except Burnout and Fallibility).

Ryan's assessment needs to be revisited in light of Gallup's Urban TPI developed specially for that context Individualized Perception, Empathy, and Input [Drive] have been carried over directly from the original TP Mission has become Commitment (with a new emphasis on working where there is the "greatest need"). In into Dedication and Caring (with a new emphasis on emotional relationships). Rapport Drive and Listening Initiator (with a new emphasis on being an advocate for students). Activation has become Stimulator (with personally dramatic). Focus and possibly Gestalt have merged to form Concept (with a new emphasis on p Innovation and Objectivity do not appear to have direct corollaries in the Urban TPI, while Involver and Pc themes specific to the Urban TPI. When compared to the ten themes in Haberman's on-line interview, the c standard and Urban TPI stand out: there is a greater emphasis on emotional relationships with students (like Students and Approach to Students themes), on positive attitudes and expectations for students (like Haber and Planning, Student Learning, and Explains Success themes), and on stamina in dealing with students (lil

and Survive in Bureaucracy themes).

A serious problem suggested by this brief overview is the lack of psychometric justification for these varie whether the constructs were created following standard instrument-design procedure and whether their outc statistically—concerns reinforced by the fact that the constructs appear to change over time (witnessed by t for them in different articles). The two companies, it should be noted, claim different roles for their intervie publications state that the purpose of his interview is to identify teachers who *could be prepared* to be succ whereas the Gallup Perceiver system is designed to identify new teachers *to be hired.*

Analysis: The TPI

Commercial teacher interviews are an important topic for educational researchers to investigate because the may not be adequately understood. The designers of commercial teacher interviews do not claim that their teaching, but rather that they identify teacher candidates who communicate the same professional values an "best" teachers. Yet, there is also a perception that the interviews are quantifiable and non-discriminatory. structured interview's greatest selling point—Gallup's Automated Teacher Screener webpage assures pote protected by regular and continuing checks for EEOC fairness for all candidates interviewed with the system" (http://education.gallup.com/attract/autoTeaScreen.asp). [3]

Furthermore, educational researchers are justified in examining the validity of commercial interviews beca are spending sizable sums of money on using them. According to their website (http://www.cobb.k12.ga.us boardagenda/Dec1098/discussion%20agenda/haberman.htm), Georgia's Cobb county school board voted i contract with Haberman Educational Foundation, Inc. for Star Teacher Selection Interview training. Cedar public their costs for using Gallup's products. According to their website, the TPI required a one-time initi per administrator plus a $500 annual fee per administrator. Shifting to the TeacherInsight interview, the dis initial consulting and system start-up costs plus $10,000 for a two-day administrator training seminar for u recurring estimated costs will be approximately $14 per candidate to complete the online screener. (http://www.window.state.tx.us /tspr/cedarhill/ch03c.htm)

Why meta-analysis is needed. The validity of the TPI has been studied in several dissertations, but not a si been published in a refereed journal. All the dissertations completed in the 1970s, when the TPI was first re significant relationships between the TPI and various criteria utilized as indicators of teaching quality. A S researcher in a commentary on the TPI concluded, "All available evidence fails to support claims made in t and by SRI sales representatives and TPI users. The users have made a premature commitment to a selectio to be accurate on the surface but fails to meet minimal requirements for instrument validity." (Haefele, 197 candidates have conducted some further validation studies since the 1970s, but the results were mixed. For synthesized a range of studies from the 1970s to the present in which quantitative relationships between the teaching quality were reported, examining the TPI's validity based on accumulated evidence.

Study selection. A portion of the studies included in our synthesis came from the collection accumulated fc et al., 2002), of which we are a part. Additional sources were found by searching ERIC and Dissertation Al "TPI" and "Teacher Perceiver Interview. " Originally, we found 25 studies that examined a relationship bet teaching quality variables. We included in our synthesis only those reporting correlation coefficients (Pears ratings and teaching-quality indicators (which we grouped as student ratings, principal ratings, classroom o gain scores). Ultimately, we included thirteen studies (all dissertations) in our synthesis; the TPI studies va design, so we chose only those with comparable measures.

Coding the studies. Pairs of coders independently coded a number of important characteristics for each stu presented in Table 1. "State" indicates where the study was conducted; "Grade" indicates grade level taugh teachers; "N" indicates the number of participating teachers; "Criterion of teaching quality" indicates outco for hiring" indicates whether the TPI data in the study were collected during the hiring process or later only of effect sizes" indicates the number of relationships between the TPI and the criterion, shown as correlatio used only the r's between the TPI total scores and the indicator of teaching quality (rather than including th and the outcome, if reported), because total score is likely most influential in the hiring process. Moreover, relationships between TPI total score and any subset of the measured outcome in addition to the criterion's the r between the TPI total score and the measured outcome's total score in order to avoid data dependence the form of percentage agreement between two coders, ranged from 84% (TPI used for hiring) to 100% (fo Discrepancies between the two coders were resolved before conducting further analyses.

Table 1. Characteristics of the 13 studies

Candidates for alternative programs may be recommended by principals or other school administrators, and sometimes are required to have met other criteria (such as having a GPA above 2.5, etc., see Guyton et al., 1991).

Alternatively certified teachers also have typically taken fewer education courses than traditionally certified teachers. However, there is also quite a lot of variation among alternative certification programs in the amount of education coursework participants actually take (e.g., Darling-Hammond, Berry & Thoreson, 2001, p.11). Because of their abbreviated nature, alternative-certification education courses are likely to cover content different from that in courses taken by traditionally certified teachers. Participants typically take their educational coursework during the on-the-job internship, though most alternative-certification programs require that participants already hold a bachelor's degree in some subject-matter field (e.g., math or history rather than education). Finally, the duration of alternative training programs is usually shorter than that of traditional education programs.

Emergency certificates may be given to even less-prepared teachers. There may not even be a "program" *per se* in which the potential teacher would enroll or participate. Nonetheless some teachers hold emergency certificates, and the requirements appear to be quite minimal.

Research Questions

In our work we are interested in making several key comparisons. For each set of comparisons we also will explore more detailed questions about possible moderator variables that may explain differences among the study results we find. Our three key research questions are:

1. Are teachers with traditional teacher certificates more or less effective at

teaching than those holding alternative teacher certificates or emergency teacher certificates? Secondarily, is there any difference in teaching effectiveness between teachers with alternative certificates and those with emergency certificates?

2. Do the effects of certification differ across the subject-matter taught, across outcomes (e.g., for teacher performance versus student achievement), and by school level (elementary versus the higher grades)?

3. Are certification effects moderated by other factors such as publication status, type of rater, and whether the levels of teaching experience of those with different certificates were controlled or adjusted in some way?

In our analyses we address the first question for the entire set of effects, then we pursue questions one and two within the three sets of certification comparisons available (traditional vs. alternative, traditional vs. emergency, and alternative vs. emergency).

Methods

The Studies

The studies in our synthesis were identified as a part of our larger synthesis project, which to date has gathered over 500 documents reporting on the relationship between teacher qualifications and the quality of teaching. In this section we describe the search process used to identify studies for the project overall, and within the project for this synthesis. We also discuss some basic descriptors of the studies in our analysis, with a focus on the kinds of outcomes examined in those studies.

Search process. The project overall aims to examine the literature on teacher qualifications (of which certification is one) and the relation of qualifications to the quality of teaching. Our overall search strategy is described in Wu et al. (2002). Five inclusion rules were used to obtain our studies:

1. Studies must have been conducted in the United States, during the years 1960 to the present.

2. Only studies examining K-12 grade teachers are included.

3. Studies must have provided sufficient data to compute an index of difference between groups of teachers with different types of certifications. Qualitative studies and case studies were not included.

4. Teaching-quality outcomes could be represented in terms of either teacher-performance measures or student-achievement outcomes.

5. Teacher-performance measures used as indicators of teaching quality should be based on
objective observations by other persons, using performance based evaluation forms or other standardized instruments. Principal ratings and student evaluations of teacher performance are
included (we found only one student evaluation study). No teacher self-reports of their own teaching outcomes were included.

For this synthesis we selected studies examining type of certification as the representation of teacher qualifications. To identify studies of different kinds of certification within our overall collection of studies, we used our EndNotes database to select studies which had examined teacher certification as the indicator of qualifications. We did a search for the words "alternat*", "emergency", and "certificat*" (the * is a wild card allowing us to identify all studies with words having the letters shown - such as "certificat" -- followed by any other letters (ion, e, etc.). Through this search we identified 39 documents. These 12 dissertations and 27 other documents were examined to see what kinds of research design were used, whether data were reported, and the like. We identified several pairs of documents that had reported on the same samples and thus were combined to represent one study. Three pairs represented dissertations and the articles that followed (Brown, 1987 and Brown et al., 1989; Hall, 1962 and 1964; Shim, 1963 and 1965). Another pair by Beery (1960, 1962) represented an article and the full research report of the same study. Also Hall and Beery used a common data set on Florida teachers and their students. After examining common data sets and excluding studies for reasons listed below, we arrived at a set of 24 studies from 21 different documents. Table

2 lists the 24 studies in chronological order and shows some basic study characteristics.

Omitted studies. Several potential studies were omitted because the data reported either were incomplete or were not commensurate with data in other studies. For example, Boser and Wiley (1988) reported principal opinions about the superiority of traditional versus alternate-certified teachers. While the data came from principals in schools where alternate-certified teachers were employed, the principals were not asked to rate the specific teachers in their schools and no comparisons of those teachers to specific traditionally certified teachers were made. By contrast, Raymond, Fletcher and Luque (2001) had compared traditionally certified

Table 2. *List of Studies and Study Characteristics*

Study	Publication Year	Publication Status	Comparison Type	Outcome
Beery	1960 (1962)	Report	Trad Prov vs. Emergency	Teacher Performance
Hall	1962	Dissertation	Trad Prov vs. Emergency	Teacher Performance
Shim	1963 (1965)	Dissertation	Trad Full vs. Emergency	Student Achievement
Bledsoe	1967	Report	Trad Prov vs. Emergency	Teacher Performance
Cornett 1	1984	Report	Trad Prov vs. Emergency	Teacher Performance
Cornett 2	1984	Report	Trad Prov vs. Emergency, Trad Full vs. Emergency	Teacher Performance
Hawk	1985	Journal	Trad Full vs. Out-of-field	Student Achievement
Brown	1987 (1989)	Dissertation	Trad Prov vs. Alternate, Trad Prov vs. Emergency, Alternate vs. Emergency	Teacher Performance
DeWalt	1987	Journal	Trad Prov vs. Emergency	Teacher Performance
Goebel	1988	Report	Trad Prov vs. Alternate	Student Achievement
Goebel	1989	Report	Trad Prov vs. Alternate	Student Achievement
Hawk	1989	Journal	Trad Prov vs. Alternate	Teacher Performance
Guyton	1991	Journal	Trad Prov vs. Alternate	Teacher Performance
Knight	1991	Journal	Trad Prov vs. Alternate	Teacher Performance
Dial 1	1992	Dissertation	Trad Prov vs. Alternate	Teacher Performance
Dial 2	1992	Dissertation	Trad Prov vs. Alternate	Student Achievement
Sandlin	1992-93	Journal	Trad Prov vs. Alternate	Teacher Performance
Jelmberg	1996	Journal	Trad Prov vs. Alternate	Teacher Performance
Miller 1	1998	Journal	Trad Prov vs. Alternate	Teacher Performance
Miller 2	1998	Journal	Trad Prov vs. Alternate	Student Achievement
Zirkel	1998	Dissertation	Trad Full vs. Out-of-field	Student Achievement

Achievement

Pezzano	1999	Dissertation	Trad Prov vs. Alternate Student Achievement	
GoldBrew	2000 (1999)	Journal	Trad Prov vs. Emergency, Trad Full vs.	
Emergency, Trad Prov vs. Out-of-field, Trad Full vs. Out-of-field Student Achievement				
Lackzo	2002 (2000, 2001)	Journal	Trad Prov vs. Alternate, Trad Prov vs. Emergency	

Student Achievement

Table 2 (Continued). *List of Studies and Study Characteristics*

Study	Publication Date	Number of Effects (Without Totals)		Number of
Samples	Number of Outcomes		Conclusion	
Beery	1960 (1962)	20 (16)	4	5 TP>E
Hall	1962	6	1	6 TP>E
Shim	1963 (1965)	6	2	3 TF<E
Bledsoe	1967	7 (4)	1	7 TP>E
Cornett 1	1984	1	1	1 TP>E
Cornett 2	1984	12 (9)	3	4 TF>E;
TF<E				
Hawk	1985	2	1	2 TF>Out
Brown	1987 (1989)	36 (15)	3	6 x 2 raters
	TP=A=E			
DeWalt	1987	12	1	12 TP>E
Goebel	1988	2	2	1 TP=A
Goebel	1989	2	2	1 TP=A
Hawk	1989	5	1	5 TP>A
Guyton	1991	1	1	1 TP<A
Knight	1991	3	1	3 TP>=A
Dial 1	1992	6 (5)	1	6 TP<=A
Dial 2	1992	18 (13)	2	16 (sample 1)
+ 2 (sample 2)		TP=A		
Sandlin	1992-93	2	1	6 + 3 TP=A
Jelmberg	1996	2	1	2 TP>A
Miller 1	1998	3 (2)	1	3 TP=A
Miller 2	1998	5	1	5 TP=A
Zirkel	1998	2	2	1 TF>Out
Pezzano	1999	3	1	3 TP=A
GoldBrew	2000 (1999)	24	4	6 (2 post x 2
times + 2 gains)		TF>E, TP>E, TF>Out, TP>Out		
Lackzo	2002 (2000, 2001)	12	4	3 TP>A,
TP>E				

teachers to others certified via the Teach for America program, but they did not report sample sizes for the separate analyses of elementary and middle-school teachers where the key results reside. Finally, some references in our collection report descriptive data concerning teachers with different types of certification, but do not report complete data on the performances of either the teachers or their students (e.g., Copley, 1974; Hutton et al., 1990; Lupone, 1961).

A third category of study that was eventually omitted also was found within the set of identified studies. That category included eight studies using designs that did not allow us to compute effects similar to those from comparison studies. Some were based

on data aggregated to the school level or higher, so had reported on such variables as the percentage of certified teachers in the school (e.g., Fetler, 1999; Mandeville & Liu, 1997). Another eight studies used regression analyses and also did not allow us to compute clean comparisons of teachers with different types of certifications (e.g., Perl, 1973).

Coding

Certification comparisons. Four main kinds of comparison appear in the data set, comparisons of traditionally certified teachers versus alternate-certified teachers, traditionally certified teachers versus those with emergency certificates, traditionally certified teachers versus out-of-field teachers (teachers fully certified in one area but teaching in a different area), and alternate-certified teachers versus emergency-certified teachers. In addition, we were also able to categorize traditionally certified teachers as having provisional (with minimal experience) or full (with more experience) certifications. This led to having seven comparisons all told. Table 3 shows the numbers of studies and effect sizes for each of the comparisons. The counts of studies do not sum to 24 because some studies provided several comparisons.

Table 3. *Numbers of Studies and Effect Sizes for Each of Seven Comparisons*

Comparison (151)	Number of studies (24)	Number of Effect Sizes
Traditional Provisional vs. Alternative	14	55
Traditional Full vs. Alternative	1	2
Traditional Provisional vs. Emergency	7	55
Traditional Full vs. Emergency	4	23
Traditional Provisional vs. Out-of-field	1	6
Traditional Full vs. Out-of-field	3	10
Alternative vs. Emergency	1	5

Indicators of teaching quality. The outcome or measure of teaching quality is another important and complex moderator variable coded in this study. Soar (1983) pointed out three types of outcome researchers have used to evaluate teacher quality: "presage variables" - including tests of teacher knowledge (e.g., the National Teacher Examination or NTE); "product variables", or tests of student achievement; and "process variables" -- classroom teacher-performance measures. The first criterion is one that we have classified in our project as an index of teacher qualifications - a competitor to certification status as an index of teacher suitability. While eventually we may examine results based on such measures, in this paper we have examined only the latter two.

We do examine measures of student achievement, which were reported in 9 of the 24 studies in our set. Soar and colleagues have argued that student achievement gains may not tell us how competent a teacher is, since many of the differences in pupil performance are attributable to influences beyond the teacher's control. Others, however, argue that student achievement is the best index of teacher quality. We coded the type of subject matter being examined by each student-achievement test, as well as information about test reliability (though we do not use that information in the current analyses).

In contrast, Soar argued that performance-based teaching-quality measures are most reliable and objective. They thus prefer to use teaching performance evaluated by performance-based tests as the indicator of teaching quality. Fifteen of our studies used teacher classroom performance as an indicator of teaching quality (or teaching effectiveness). We coded the subject matter being taught (when teacher-performance measures were reported), as well as information about who had made the ratings of teacher behaviors (raters from the teacher's school, such as a principal; outside raters, such as a superintendent, the experimenter, or state officials; or students). Some instruments also included subscales that resembled personality measures (e.g., Beery, 1960; Bledsoe et al., 1967). We thus differentiated these personality-like measures from other teacher-performance scores.

Finally studies often reported both subtest scores and total scores for teaching performance and student achievement, thus many of our studies have multiple outcomes. As a first step in our coding, we coded results for every available subtest score and total score for each study. To avoid dependence among the outcomes, both total scores and subtest scores from the same study should not be analyzed together. As a partial way to address this dependence, we eliminated total scores when both totals and subtest scores were presented, and created five sets of outcomes, based on student-achievement total scores, student-achievement subtest scores, teacher-performance total scores and teacher-performance subtest scores and teacher personality-like measures. Then for another set of analyses we omitted the subtests and examined only the total scores. Differences among these student achievement and teacher measures are thus of great interest in our analyses.

Other coded variables. Several additional variables were coded, including the

school level at which the teachers were employed (elementary, secondary or mixed), the state in which the study was conducted, the publication date and type (dissertation, journal article or report), and whether the study had controlled for differences in prior teaching experience among the certification groups. Table 4 shows the variables and the counts of studies and effect sizes for each variable. Also Table A1 in our appendix shows the counts of studies and effect sizes for each variable within each comparison group.

Table 4. *Coded Variables*

Variable	Levels of the Variable	Number of Studies	Effects
Outcome	Student achievement total Student achievement subtest Teacher performance total Teacher performance subtest Personality-like measure	8 3 8 11 6	62 17 18 58 37
Subject	Math Science Reading Language Music Other	8 3 5 4 1 13	31 13 14 11 2 66
School level	Elementary Secondary Mixed	11 5 9	102 51 39
Experience controlled	Experience controlled Controlled	10 15	not 78 114
Publication type	Dissertation Journal article Report	7 11 6	77 59 56
Type of rater	Inside observer Outside observer Student evaluation No rater	6 7 1 10	42 79 3 68

Effect Sizes

The index used to represent differences among the certification groups is the standardized mean difference, often called Cohen's d or Glass's effect size. Specifically we computed

$$g = (\bar{Y}_1 - \bar{Y}_2)/S_p,$$

where $\bar{Y}_1$ is the mean for one group of teachers, $\bar{Y}_2$ is the mean for the comparison group,

and S_p is the standard deviation pooled across the two groups. For comparisons involving traditionally certified teachers, effect sizes were computed so that positive values represent the superiority of the traditionally certified teachers or their students (i.e., $\overline{Y}_1$ is the mean for traditionally certified teachers). For the contrast comparing alternate versus emergency certifications, positive values indicate that alternately certified teachers outperformed those with emergency certificates.

In a few cases data were not available to compute the standard formula above. Standard translations from t and F values were used, and in one case where data had been dichotomized formulas from Haddock et al. (1998) were applied.

Effect sizes were corrected for small-sample bias using Hedges's correction, $d=g*c(m)$ where $m = n_1+n_2 - 2$ and the unbiased effect size d was weighted in analyses by the inverse of its variance $Var(d)= ((n_1+n_2)/(n_1*n_2))+(d^2/(2*(n_1+n_2)))$.

Analyses

Our analyses follow methods outlined in the *Handbook of Research Synthesis*. We use fixed-effects (Hedges, 1994) and random-effects (Raudenbush, 1994) categorical models for most of the investigations. One important statistic used in this approach is the homoegeneity test, denoted below as Q(df). There are several forms of Q, each being a weighted variance. Under the appropriate null hypothesis for each Q, the statistic follows a chi-square distribution with degrees of freedom (df) that relate to the numbers of effects or groups being compared in each analysis. Related to Q is Birge's ratio, Q/df (Birge, 1932). When the null hypothesis for each Q is true, the expected value of Q is its degrees of freedom. Thus a Birge ratio near 1 indicates agreement with the null model. Ratios

much larger than 1 typically reflect inconsistencies among study results or very large between-groups differences.

Results

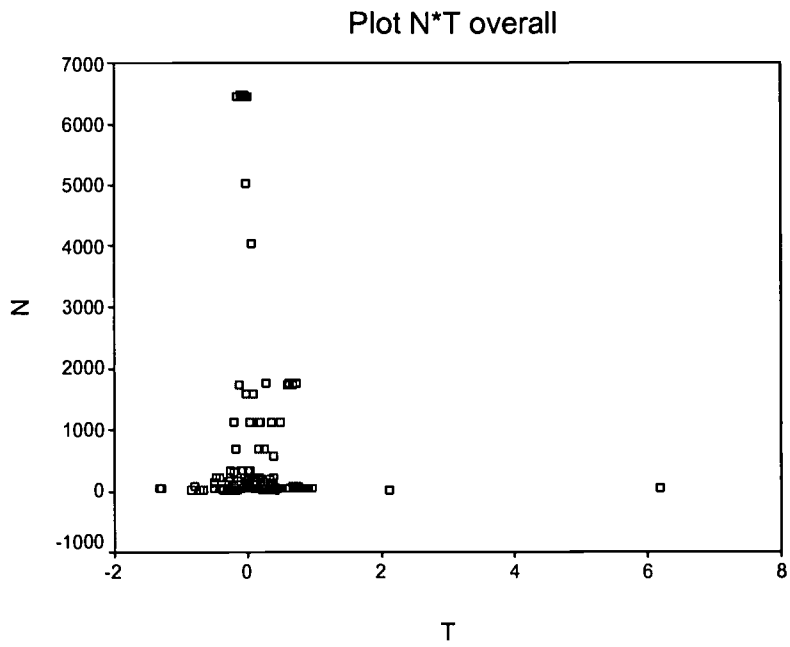
Description of the Studies

The 24 studies in our data set allowed for the computation of 192 effect sizes. Studies often had multiple samples (thus provided multiple comparisons) and each sample may have been measured on several outcomes. When multiple samples were examined, we used the most fine-grained subsets for our computations. The number of samples per study ranged from 1 to 4 and the number of outcomes ranged from 1 to 16. The largest number of effects was obtained from Brown (1987, 1989), which produced 36 effect sizes (for 3 samples x 6 outcomes x 2 raters). In most of our further analyses we eliminated effects representing total scores when studies had reported both total and subtest scores, reducing the number of effects to 153. For some analyses we examined total scores instead of subtests - this data set included 109 effects.

Seven kinds of comparisons were found amongst the 24 studies, with nearly half of the comparisons (94 of 192), representing traditional- versus emergency-certified teachers. Just over a third of the effects (70 of 192) were comparisons of alternate- and traditionally-certified teachers, and about 10 percent of the effects compared traditional with out-of-field teachers or alternate- versus emergency-certified teachers. Clearly also most of the comparisons involved new teachers; the three sets of comparisons involving traditional provisional teachers (148 effects) made up just over three-fourths of the effects in our data. Almost 60 percent of the effects (113 of 192) represented teacher outcomes.

Assessing the Presence of Publication Bias

Publication bias is the tendency for a set of empirical results to be biased because results that have not reached statistical significance are not published. Publication bias should be less of a problem when unpublished documents, such as dissertations and reports, are included in a review. One way to assess whether publication bias is likely to be an issue for a set of studies is to examine a funnel plot. Since effect sizes from small studies will typically show more variability among those from larger studies (and since there will typically be fewer of the latter), a funnel plot of sample size versus estimated effect size should look like a funnel if there is no publication bias. The following funnel plot for the effects from our 24 studies is fairly symmetric and it has a funnel shape except for one effect size bigger than 6. This plot shows no apparent publication bias in this meta-analysis. This is not too surprising considering that over half of our studies were dissertations or unpublished reports.



Overall Homogeneity Test

A homogeneity test of all effects using the fixed-effects model was first conducted. The studies do not all appear to arise from a single population with a common effect size ($Q(191) = 641.68$, $p<.0001$). This is not unexpected since there is a great deal

of diversity in the set of effects and several different kinds of contrasts are included. Therefore, a random-effects model is more appropriate to describe the average effect across all studies. Under the random-effects model, the estimated common effect size is 0.08 standard deviations, which is not significantly different from zero with a 95% confidence interval (CI) ranging from 0.00 to 0.17 and a standard error of 0.044.

A quick inspection of the data shows that two effects, from Zirkle (1988), are very large relative to the rest. These values ($d = 2.1$ and $d = 6.1$) were the only effects representing music outcomes, and were based on a comparison of traditional fully certified versus out-of-field teachers. The effects were considered to be outliers and were eliminated from further analysis. The random-effects mean recomputed without the two Zirkle effect sizes was slightly lower, with a value of 0.06 (SE=0.029), and the smaller standard error has led to a narrower confidence interval; the 95% CI covers from 0.00 to 0.12.

Because so many studies allowed us to compute both subtest and total-score effect sizes, the set of 190 effects suffers from serious dependence issues. To reduce the influence of studies that had reported both totals and subtests, we eliminated 39 duplicative total scores from the data set. The remaining 151 effects were still significantly heterogeneous ($Q(150) = 458.64$, $p < .0001$). The random-effects mean based on the final set of 151 effects is 0.08 (SE= 0.029) with a 95% CI of 0.03 to 0.14. The most conservative between-studies variance estimated for the set of 151 effects was 0.079, equivalent to a standard deviation (SD) of 0.28 - just over a fourth of a standard-deviation unit. We can interpret this value by supposing that the population of effect sizes is normal, with a mean equal to the random-effects mean (0.08). If the distribution of true

effects is centered on this mean, and has SD=0.28, then roughly 61 percent of all effects will be positive, showing superiority of traditional certification and advantage of alternate routes over emergency certification. However, this collection of effects is very diverse and so it is not surprising to find a wide range of values.

Certification-type Comparisons

Since the purpose of this study is to detect the effect of different certificate types on the quality of teaching, we are most interested in estimating the population effect sizes for each comparison type. We also ask whether the population effect sizes for each comparison group differ from each other, and whether the groups of comparisons are internally homogeneous. When they are not internally homogeneous, we investigate what variables explain the variation in effect sizes within each comparison group. Hence, our first analysis is an analogue to the analysis of variance of the 151 effects using *Comparison type* as the factor.

Preliminary to the analysis we examined whether effects for the two sorts of traditionally certified teachers could be considered together in comparisons with the nonstandard routes. In two cases (comparisons with alternate and emergency certification) the results for the two traditional subgroups are very similar. No significant differences were found between the effects for the provisional and fully-certified groups versus the alternate-route teachers (z=0.14), or versus the emergency certified teachers (z=0.82). Therefore, studies of these two types of comparisons are merged together into "traditional certificate versus alternate" and "traditional certificate versus emergency" groups. By contrast, the two sets of effects examining out-of-field teachers were significantly different on average (z=2.87) and thus were not combined. Appendix Table A2 shows the

analysis of the seven groups.

Table 5 shows the between-groups variance explained by *Comparison type*, and important statistics within each comparison group, for the fixed-effects analysis.

Table 5. *Fixed-Effects Categorical Model by Comparison Type*

	Q	df	p-value	Birge Ratio		
	Mean Effect		SE of the Mean			
Total	458.66*	150	<.0001	3.06	--	
	--					
Between groups	104.65*	4	<.0001	26.16	--	
	--					
Within groups	353.98*	146	<.0001	2.42	--	
	--					
Traditional vs. Alternate	151.06*		55	<.0001	2.75	-
0.05*	0.007					
Traditional vs. Emergency	174.45*	75	<.0001	2.33	0.11*	
	0.025					
Trad provisional vs. Out-of-field	7.40	5	.19	1.23	0.03	
	0.114					
Traditional full vs. Out-of-field	13.68	7	.06	1.95	0.39*	
	0.053					
Alternate vs. Emergency	7.39	4	.12	1.85	-0.37*	
	0.171					

Table 5 reveals that the sets of effects for out-of-field comparisons and for the comparison of alternate-certified and emergency teachers are homogenous, indicating that all of the effects for each of these contrasts show similar results and likely arise from the same population. We return to these results below. Unexplained variation is found within the two other groups (traditional versus alternate, and versus emergency), particularly for the effects involving alternate certification. The Birge ratio for this set of effects indicates that these effects are 2.75 times more variable than would be expected due to random variation. Further analyses explored other factors that might explain the differences

among the effect sizes within these two groups.

Traditional versus Alternate Certification

Seven different study characteristics were examined to see if they related to the sizes of the standardized mean differences between traditional and alternately certified teachers. The seven characteristics analyzed did relate to the magnitudes of these effects. Within the traditional versus alternate-certification comparisons, every potential predictor variable explained some of the variation in the 56 effect sizes, but no predictor accounted for all of the variation. The Birge ratios for between-groups variation reflect the strength of these predictors. *School level* had the biggest Birge ratio of 7.91, and variables *Subject* (Birge=7.65), *Experience controlled* (6.14), and *State* (5.32) all had ratios bigger than 5. *Publication type* (3.52) and *Rater* (2.99) were less predictive. Table 6 shows the results of these categorical analyses. Means (and 95% confidence limits) for effects that differed significantly from zero are shown in bold.

Many of the sets of means do not seem to follow any clearly explainable pattern, and nearly all show considerable heterogeneity within the subgroups examined. One exception is that the results by *State* show consistency within 5 of the 7 states studied. Also very strong differences exist between states, with two states (Arizona and New Hampshire) showing strong superiority for traditionally certified teachers and two others (California and Texas) showing advantages for alternate-route teachers. This predictor seems particularly important because states can have very different certification rules for both traditional and alternate routes. A further step in our investigation will be to examine the requirements for these states to see whether particular differences in requirements can be identified.

Because none of the potential explanatory variables fully accounted for variation in the traditional-versus-alternate-certification comparisons, we estimated the mean effect for these studies under the random-effects model. The random-effects model assumes there is a population of effects varying randomly around an average true effect. An estimate of this variation (or uncertainty) is incorporated into the mean and its standard error to allow for this spread in true effects. The random-effects analysis showed a between-studies standard deviation of 0.17, and a

Table 6. *Categorical Analyses for Traditional versus Alternate Certification Groups*

	k	Q	p	Lower Limit Ratio	Mean Effect	Upper Limit	SE	Birge
Between Outcomes		14.18	.007					3.55
Within Outcomes		136.87	<.0001					
Stu Subtest	11		33.91	.0002	**-0.08**	**-0.06**	**-0.05**	0.009
								3.39
Stu Total	17		26.67	.045	-0.05	-0.02	0.01	0.014
								1.67
Tch Subtest	22		72.35	.00000	-0.03	0.03	0.10	0.033
								3.45
Tch Total	1		0.00	.	-0.94	-0.17	0.61	0.398
	.							
Tch Persnlty	5		3.95	.413	-0.10	0.10	0.30	0.103
								0.99
Between Rater Types		8.98	.029					
								2.99
Within Rater Types		142.08	<.0001					
Inside	11		27.04	.0026	-0.09	0.04	0.18	0.067
								2.70
Outside	14		29.35	.0058	-0.15	-0.03	0.09	0.059
								2.26
Student	3	18.77	<.0001	-0.02	0.07	0.16	0.045	
								9.38
None	28		66.92	.00003	**-0.07**	**-0.05**	**-0.04**	0.007
								2.48
Between School Levels		15.81	.0004					
								7.91
Within School Levels		135.25	<.0001					
Elementary	34	84.90	<.0001	**-0.07**	**-0.05**	**-0.04**	0.007	
								2.57
Secondary	2	0.04	.84	-0.09	0.22	0.53	0.157	
								0.04
Mixed	20	50.31	.00012	**0.01**	**0.08**	**0.15**	0.036	
								2.65
Between Subjects		30.61	<.0001					
								7.65
Within Subjects		120.44	<.0001					
Math	10		24.95	.003	**-0.13**	**-0.10**	**-0.08**	0.013
								2.77
Science	1		0.00	.	**-0.12**	**-0.07**	**-0.02**	0.025
	.							
Reading	7		4.47	.61	-0.05	-0.03	0.00	0.014
								0.74
Language	4		2.22	.53	-0.06	-0.02	0.01	0.018
								0.74
Other	34		88.81	<.0001	-0.03	-0.00	0.03	0.016

	2.69							
Between Exp Ctrl		6.81	.009					
	6.81							
Within Exp Ctrl		144.25	<.0001					
Ctrl	41		95.80	<.0001	**-0.06**	**-0.05**	**-0.04**	0.008
	2.39							
Not	15		48.44	<.0001	-0.02	0.05	0.12	0.037
	3.46							
Between Pub Types		3.18	.204					
	1.59							
Within Pub Types		147.88	<.0001					
Dissertation	26		67.84	<.0001	**-0.06**	**-0.05**	**-0.03**	0.008
	2.71							
Journal Article	20		69.37	<.0001	-0.05	0.00	0.06	0.030
	3.65							
Report	10		10.67	.30	-0.16	-0.02	0.12	0.072
	1.19							
Between States		31.90	<.0001					
	5.32							
Within States		119.16	<.0001					
Arizona	6	0.17	.99	**0.09**	**0.38**	**0.66**	0.150	
	0.03							
California	2	0.33	.57	**-0.76**	**-0.50**	**-0.25**	0.129	
	0.33							
Georgia	8	8.05	.33	-0.15	-0.06	0.03	0.046	
	1.15							
N. Hampshire	2	0.03	.87	**0.07**	**0.36**	**0.65**	0.149	
	0.03							
New Jersey	3	1.99	.37	-0.33	-0.14	0.05	0.096	
	0.99							
N. Carolina	5		15.51	<.0001	-0.09	0.17	0.42	0.129
	3.88							
Texas	30	93.08	<.0001	**-0.06**	**-0.05**	**-0.03**	0.007	
	3.21							

Table 7. *Analyses of Total Scores for Traditional versus Alternate Certification Groups*

	k	Q	Mean effect	SE of the Mean	
	95% CI Lower limit		95% CI Upper limit		
Between Outcomes Within Outcomes			7.18** 98.74**		
Student achievement	23	47.64**	-0.05*	0.009	-0.06	-0.03
Teacher outcomes	17	51.10**	0.05	0.036	-0.02	0.12
Between Rater Types Within Rater Types				1.74 49.35**		
Inside	11	19.33*	0.07	0.071	-0.07	0.21
Outside	3	11.26*	-0.09	0.112	-0.31	0.13
Student	3	18.77*	0.07	0.041	-0.02	0.16
Between School Levels Within School Levels				12.75** 93.16**		
Elementary	18	41.29**	-0.05	0.009	-0.06	-0.03
Secondary	1	0.00	0.49*	0.224	0.01	0.89
Mixed	21	51.87**	0.06	0.036	-0.01	0.13
Between Subjects Within Subjects			10.08* 95.84**			
Math	5	20.11**	-0.10*	0.022	-0.13	-0.05
Science	1	0.00	-0.07*	0.025	-0.12	-0.02
Reading	10	12.19	-0.03	0.022	-0.07	0.01
Language	3	2.02	-0.00	0.025	-0.05	0.05
Other	21	61.52**	-0.03*	0.012	-0.05	-0.01
Between Exp Control Within Exp Control				5.63* 100.27**		
Controlled	31	66.77**	-0.04*	0.009	-0.06	-0.03
Not controlled	9	33.50**	0.05	0.040	-0.03	0.13
Between Pub Types Within Pub Types			13.65** 92.26**			

Dissertation	21	50.14**	0.08*	0.034	0.01	-
0.15						
Journal Article	15	40.99**	-0.05*	0.009	-0.07	-
0.03						
Report	4	1.14	-0.14	0.082	-.30	
	0.02					
Between States Within States		26.54** 79.37**				
Arizona	6	0.17	0.38*	0.151	0.09	
	0.68					
California	1	0.00	-0.51*	0.183	-0.87	-
0.15						
Georgia	4	4.89	0.02	0.071	-0.12	
	0.16					
New Hampshire	2	0.03	0.36*	0.149	0.07	
	0.65					
New Jersey	3	1.99	-0.14	0.096	-0.33	
	0.05					
North Carolina	5	6.47	0.25	0.169	-0.09	
	0.58					
Texas	19	65.83**	-0.04*	0.009	-0.06	-
0.03						

mean of -0.01. The 95 percent confidence interval for the mean ranged from -0.07 to 0.05, showing that the mean of the population effects is essentially zero and that the effects are roughly evenly split between positive effects (favoring traditional teachers) and negative ones (favoring those with alternate certification). If the effects follow a normal distribution, ninety-five percent of the population effects are expected to fall between -0.33 and 0.33. Even the largest of these effects are not strong in either direction.

Total-score analysis. The analyses just described included effects from all possible subtests when both subtests and total scores were presented. Those analyses give more detail on the scores from varying kinds of subtests, but they also exhibit considerable dependence because often several subtests from the same subjects are analyzed together. We present additional analyses using total scores (omitting subtests) when a study had presented both. Comparisons of traditional and alternate-route teachers were represented by 40 effects. Table 7 shows analyses of the seven predictors on the set of total-score effects. Again most of the predictors explained significant amounts of variation, but did not fully account for between-studies differences.

Traditional versus Emergency Certification

The same seven study characteristics were used to analyze the effects for the traditional versus emergency comparisons. For these 76 effects, only three predictors explained significant amounts of variation - *State, Publication type* and *Outcome.* However, none of these account for all of the variation among the effects (see Table 8). The variable *State* has the biggest Birge ratio of 9.13, and again some dramatic between-state differences appeared. Arizona and Florida showed significant advantages for traditionally certified teachers. Only Maryland showed emergency teachers outperforming

traditional teachers. The Maryland results all arose from Shim (1963) and represented a rather unusual comparison - Shim compared students who had been taught by a *series* of four teachers all of whom were either traditionally certified or emergency-certified.

Again because none of the potential explanatory variables fully accounted for variation in the traditional-versus-emergency-certification comparisons, we estimated the random-effects mean. The random-effects analysis showed a between-studies standard deviation of 0.32, and a mean of 0.14. The 95 percent confidence interval ranged from 0.05 to 0.23, showing that the mean of the population effects is positive and that the effects, on average, favor traditional teachers. However these effects show much greater spread than the alternate-route comparisons, and if the effects have a normal distribution centered on 0.14, we would expect that two-thirds of the population effects comparing traditional and emergency teachers would be positive (favoring traditionally certified teachers) and only a third negative. Ninety-five percent of the population effects would fall between -0.49 and 0.77.

Total-score analysis. Again we analyzed the set of effects computed for total scores, omitting subtests when both totals and subtests were available. Fifty-two effects were available for the traditional- versus emergency-certification comparison. Table 9 shows that the same predictors explained significant between-studies differences for the totals that were significant for the subtest effects, but again no predictor fully accounted for differences in the effects.

Traditional Certification versus Out-of-field Teaching

The comparison of traditionally certified teachers versus out-of-field teachers differs somewhat from the other comparisons we have made, because often teachers who

are classified as out-of-field hold traditional teaching certificates. However, they are teaching in an area for which their certification did not prepare them.

Table 8. *Categorical Analyses for Traditional versus Emergency Certification Groups*

	k	Q	p	Lower Limit	Mean Effect	Upper Limit	SE
	Birge Ratio						
Between Outcomes		16.97	.002				
3.39							
Within Outcomes		157.48	<.0001				
Stu Subtest	6	4.80	.44	**0.14**	**0.41**	**0.67**	0.135
0.96							
Stu Total	24	52.56	.0004	**0.07**	**0.17**	**0.26**	0.047
2.29							
Tch Subtest	20	41.49	.00207	-0.08	-0.00	0.07	0.039
2.18							
Tch Total	3	2.67	.26	-0.16	0.14	0.44	0.154
1.34							
Tch Persnlty	23	55.95	.00009	**0.09**	**0.19**	**0.30**	0.052
2.54							
Between Rater Type		5.69	.06				
2.85							
Within Rater Type		168.77	<.0001				
Inside	9	7.64	.47	-0.16	0.01	0.17	0.085
0.95							
Outside	37	100.96	<.0001	**0.02**	**0.08**	**0.14**	0.033
2.80							
None	30	60.16	.0006	**0.10**	**0.19**	**0.28**	0.044
2.07							
Between School Levels		1.19	.55				
0.59							
Within School Levels		173.27	<.0001				
Elementary	29	76.58	<.0001	**0.03**	**0.12**	**0.22**	0.048
2.74							
Secondary	32	82.66	<.0001	**0.02**	**0.09**	**0.15**	0.035
2.67							
Mixed	15	14.02	.45	**0.04**	**0.15**	**0.27**	0.057
1.00							
Between Subjects		5.87	.21				
1.47							
Within Subjects		168.59	<.0001				
Math	12	29.06	.002	**0.07**	**0.20**	**0.33**	0.066
2.64							
Science	6	10.97	.05	-0.06	0.13	0.32	0.098
2.19							
Reading	6	13.02	.02	**0.05**	**0.26**	**0.47**	0.107
2.60							
Language	6	6.26	.28	-0.03	0.18	0.39	0.107
1.25							
Other	46	109.27	<.0001	**0.01**	**0.07**	**0.13**	0.031
2.43							

Between Control Groups 0.42		0.42	.52				
Within Control Groups			174.03	<.0001			
Ctrl 2.09	53	108.81	<.0001	**0.06**	**0.12**	**0.17**	0.028
Not 2.96	23	65.22	<.0001	-0.03	0.08	0.19	0.055
Between Pub. Types 3.79		7.59	.023				
Within Pub. Types		166.86	<.0001				
Dissertation 2.58	17	41.23	.0005	-0.17	0.00	0.17	0.085
Journal Article 3.15	24	72.41	<.0001	-0.01	0.06	0.13	0.036
Report 1.57	35	53.22	.02	**0.11**	**0.19**	**0.26**	0.038
Between States 9.13		63.94	<.0001				
Within States		110.51	.0008				
Arizona 0.17	6	0.86	.97	**0.12**	**0.28**	**0.43**	0.077
Florida 0.92	22	19.28	.57	**0.36**	**0.50**	**0.64**	0.071
Georgia 0.82	5	3.27	.51	-0.09	0.03	0.16	0.065
Maryland 0.29	6	1.47	.92	**-0.84**	**-0.52**	**-0.20**	0.163
National 2.78	12	30.56	.001	**0.07**	**0.19**	**0.32**	0.064
Norht Carolina 0.56	8	3.92	.79	-0.21	-0.04	0.13	0.089
Texas 3.96	5	15.83	.003	-0.35	-0.06	0.23	0.148
Virginia 3.21	12	35.33	.0002	-0.09	-0.00	0.08	0.044

Table 9. *Analyses of Total Scores for Traditional versus Emergency Certification Groups*

	k	Q	Mean effect	SE of the Mean	95% CI Lower limit	95% CI Upper limit
Between Outcomes Within Outcomes			5.77* 124.95**			
Student achievement	30	60.16**	0.19*	0.044	0.10	0.28
Teacher outcomes	22	64.79**	0.05	0.039	-0.03	0.13
Between Rater Types Within Rater Types					0.85	51.27**
Inside	9	22.84**	0.10	0.088	-0.07	0.27
Outside	12	14.48	0.12	0.079	-0.04	0.27
Student	3	13.95**	0.01	0.089	-0.16	0.19
Between School Levels Within School Levels				7.65* 123.07**		
Elementary	16	40.71**	0.10	0.076	-0.05	0.24
Secondary	26	74.91**	0.07	0.036	0.00	0.14
Mixed	10	7.45	0.28*	0.067	0.15	0.41
Between Subjects Within Subjects					6.44	124.29
**						
Math	12	29.06*	0.20*	0.066	0.07	0.33
Science	6	10.97	0.13	0.098	-0.06	0.32
Reading	6	11.52*	0.24*	0.107	0.03	0.45
Language	6	7.95	0.20	0.107	-0.01	0.41
Other	22	64.79**	0.05	0.039	-0.03	0.13
Between Exp Control Within Exp Control				0.14 130.59**		
Controlled	31	73.17**	0.10*	0.034	0.04	0.17
Not controlled	21	57.42**	0.13*	0.057	0.02	0.24
Between Pub Types Within Pub Types			4.63 126.10**			
Dissertation	30	79.74**	0.10*	0.033	0.04	0.16
Journal Article	13	25.92*	0.00	0.099	-0.19	0.20
Report	9	20.44**	0.26*	0.083	0.10	0.42
Between States Within States		26.54** 79.37**				
Arizona	6	0.86	0.28*	0.077	0.12	0.43
Florida	10	10.67	0.50*	0.105	0.30	0.71
Georgia	3	7.89*	0.19	0.113	-0.03	0.41
Maryland	6	1.49	-0.52*	0.163	-0.84	-0.20
National	12	30.56**	0.19*	0.064	0.07	0.32
North Carolina	2	0.00	0.00	0.177	-0.35	0.35
Texas	1	0.00	-0.27	0.327	-0.91	0.37
Virginia	12	35.33**	-0.01	0.044	-0.09	0.08

The analyses of the effects comparing fully or provisionally traditional-certified teachers to out-of-field teachers show contrasting results. Table 5 shows that for provisional teachers (those with few years of experience) there is virtually no difference between those teaching in-field and out-of-field. Six effects produced a mean of 0.03 standard-deviation units and a confidence interval that ranged from -0.19 to 0.25. For teachers with full certification, a notable difference is found. The mean effect (shared by all eight effects) was 0.39 - over a third of a standard-deviation unit's difference. The 95% confidence interval for this mean ranged from 0.29 to 0.49, and differed significantly from zero. The key difference between the two sets of results is that teachers with full certification have several years of teaching experience, presumably often in the area in which their certification was earned. When a teacher is assigned to teach out of his or her area of expertise, lack of experience in the new area (particularly as compared to experienced teachers teaching in-field) seems to have considerable impact. Of course there may be other potential explanations if unreported variables confounded with the assignment out-of-field also exist (e.g., if the "best" teachers are kept in their area of certification and poorer teachers are more likely to be assigned out-of-field).

Alternate versus Emergency Certification

The final comparison is the least substantiated of the five comparisons in our set. Only five effects, all from a single study (Brown, 1987) represent this comparison. According to Brown's findings, teachers with emergency certificates outperformed alternate-route teachers on all but one teacher outcome. In fact, one of the results from Brown's data showed an advantage of more than a full standard-deviation unit on one of the teacher personality measures (d=1.37 on *teacher growth and responsiveness*). While

we do not want to greatly emphasize these results from one study, the findings are consistent with the idea that more extensive training does not always lead to better teaching. Also they hint at one critical issue not dealt with by most of our studies - different kinds of individuals may end up pursuing the various routes to teaching. Thus other uncontrolled factors may differ between these groups.

Discussion

Traditional versus Alternate Certification

This meta-analysis study based on 24 studies found that overall, traditionally certified teachers and alternatively certified teachers perform equivalently. The magnitude of the difference varied by the location (state), type of outcome, school level, subject taught, whether teacher experience was controlled, and the type of rater, but even within each sort of study results did not usually agree fully. Teachers with traditional certificates tended to outperform teachers with alternative certificates in some states, but not in others. Dissertations tended to favor alternately certified teachers but journals and reports showed virtually no differences.

Because our results revealed quite a bit of variation in the differences between teachers holding these two kinds of certificates, we eventually will explore several other potential predictors of the differences. For instance, finer-grained classifications of the outcomes that have been examined may be useful. Our results suggests that although teachers from alternative training program are generally trained for less time than teachers with traditional certificates, by the end of their training programs, their outcomes appear to be similar to those for traditionally certified teachers.

Traditional versus Emergency Certification

In our meta-analysis, traditionally certified teachers generally outperformed emergency certified teachers. Emergency certified teachers are the least prepared in terms of taking educational courses and experiencing student teaching. Our results support that a certain amount of educational coursework and training on teaching skills improves the quality of teaching outcomes. However, we must be cautious because studies reported little explicit information about the levels of training of the two groups of teachers. Also considerable variation was found in the sizes of the effects from comparisons of traditional and emergency certified teachers. Indeed, about one-third of the true effects that could underlie the observed results we have collected would show negative results - favoring emergency certified teachers.

Out-of-field Teaching

The results for comparisons with out-of-field teachers are mixed. Comparisons of new (provisional) traditionally certified teachers, teaching in-field, with certified out-of-field teachers show no significant differences. Perhaps new teachers, regardless of their areas of training, will look similar in whatever field they are assigned to teach. However, teachers with full traditional certificates appear to significantly outperform out-of-field teachers. This result may reflect the role of subject-matter knowledge in improving teaching quality, or perhaps reflects the role of experience teaching in a field. This is difficult to ascertain, however, because within most of the studies in this category, teachers varied in their levels of experience. Considering the extent of out-of-field teaching that has been reported by Ingersoll (1999), the number of studies in our set provide meager data on the full range of situations in which out-of-field instruction is occurring.

Alternate versus Emergency

All of the effect sizes for the comparison of alternative and emergency teachers arose from one study (Brown, 1987). In Brown's study, emergency certified teachers had more teaching experience than alternatively certified teachers, and the emergency certified teachers outperformed alternatively certified teachers. However, this result may not be very generalizable and is based on fewer than 30 teachers. Again here the amount of information that is available to make strong generalizations is limited.

Limitations

This study suggested that the effects of different teacher certificates vary significantly between measures of student achievement and teacher classroom performance and "personality". However, our analyses were based on rather gross classifications of the outcomes available. In our future work we will examine finer-grained classifications of the outcomes that were measured, as well as investigating the psychometric characteristics of the instruments used (reliability, whether standardized instruments were used, etc.).

Our future work will also explore additional design and substantive study characteristics that may relate to the sizes of differences. Specifically we plan to examine whether other pre-existing differences (other than teaching experience) were controlled when comparisons were made, and whether different sampling designs led to differences in outcomes. While it would have been useful to have information on the ages of the teachers in the comparison groups, this was also typically not reported. Also it would also be informative to characterize exactly what levels of training the teachers studied in these different investigations had received; that was not possible given the data available in the

reports. We hope to be able to pursue some information about state-level program requirements from other documents (e.g., NASDTEC, 2002). It is not yet clear whether we will be able to connect that to our studies in a way that will be valid and informative.

Because almost all the studies provided multiple outcome measures for teaching quality, dependence among effect sizes was a problem in our anlayses. We reduced the dependence by eliminating total scores and by grouping effects by outcome type. The former approach effectively deals with dependence but unfortunately resulted in the loss of information. Future analyses on finer-grained subsets of effects should reduce further the issue of dependence.

Finally our analyses did not include all of the studies that have been done to examine the question of differences in certification. Several studies using regression-analysis design were omitted because comensurate effect indeices could not be computed from those studies. Also case studies were not included, and we suspect a richer set of data may be available from those more-detailed investigations. Our future work aims to incorporate some information from those additional studies.

Conclusion

Our findings imply that traditional teacher training is at least as effective as alternate-route training and more effective than minimal (emergency) certification. However, clearly some alternative teacher-training programs are equally effective in providing quality teachers, and one important predictor of differences in program effectiveness was the location where teachers were studied (and often trained). The role of experience was highlighted in our comparisons of in-field and out-of-field teachers, where differences were not apparent for new teachers but strongly favored more

experienced in-field teachers.

An additional overarching finding is that the studies of these alternate routes to teaching are highly various and not always well reported. Multiple confounded study characteristics appear to relate to the magnitudes of differences that were found. Yet in addition much information that would have been of use to our analyses was not reported. Our last statement is an appeal to future authors in this area - please report information as fully as possible to promote the future use of your findings and the eventual cumulation of knowledge about this important issue for educational policy.

References

*Beery, J. R. (1960). *Professional preparation and effectiveness of beginning teachers.* Coral Gables, FL: Graphic Arts Press, University of Miami.

*Beery, J. R. (1962). Does professional preparation make a difference? *The Journal of Teacher Education, 13*(4), 386-395.

Birge, R.T. (1932). The calculation of errors by the method of least squares. *Physical Review*, 40, 207-227.

*Bledsoe, C., Cox, J.V., & Burnham, R. (1967). *Comparison between selected characteristics and performance of provisionally and professionally certified beginning teachers in Georgia*. Unpublished report. (ERIC Document No. ED015553).

Boser, J.A., & Wiley, P.D. (1988). An alternative teacher preparation program: Is the promise fulfilled? *Peabody Journal of Education, 65*(2), 130-142.

*Brown, D.N. (1987). *The effectiveness of alternative teacher certification programs.* Unpublished doctoral dissertation, New Mexico State University.

*Brown, D., Edington, E., Spencer, D.A., & Tinafero, J. (1989). A comparison of alternative certification, traditionally trained, and emergency permit teachers. *Teacher Education and Practice, 5*(2), 21-23.

Copley, P. O. (1974). *A study of the effect of professional education courses on beginning teachers*. Springfield: Southwest Missouri State University (ERIC Document No. ED098147).

*Cornett, L. M. (1984). *A comparison of teacher certification test scores and performance evaluations for graduates in teacher education in arts and sciences in three southern states.* Atlanta GA: Southern Regional Education Board.

Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis*, *23*(1).

*Dewalt, M., & Ball D. W. (1987). Some effects of training on the competence of beginning teachers. *Journal of Educational Research, 80*(6), 343-347.

*Dial, M. (1992). *An evaluation of an alternative teacher certification program (teacher certification, Texas).* Unpublished doctoral dissertation (Ed.D.), University of Houston.

Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives 7.*

*Goebel, S.D., Ronacher, K., & Sanchez, K.S. (1988). Alternative Certification Program

1987-1988. Houston Independent School District, TX.

*Goebel, S.D., Ronacher, K., & Sanchez, K.S. (1989). Alternative Certification Program 1988-1989. Houston Independent School District, TX.

*Goldhaber, D.D., & Brewer, D.J. (1999). Teacher licensing and student achievement. In M. Kanstoroom & C.E. Finn (Eds.). *Better teachers, better schools*. Washington, DC: Thomas B. Fordham Foundation (pp. 83-102).

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

*Guyton, E., Fox, M., & Sisk, K. (1991). Comparison of teaching attitudes, teacher efficacy, and teacher performance of first year teachers prepared by alternative and traditional teacher education programs. *Action in Teacher Education, 13*(2), 1-9.

Haddock, C.K., Rindskopf, D., & Shadish, W.R. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. *Psychological Methods, 3*, 339-353.

*Hall, H.O. (1962). *Effectiveness of fully certified and provisionally certified first-year teachers in teaching certain fundamental skills*. Unpublished doctoral dissertation, University of Florida.

*Hall, H.O. (1964). Professional preparation and teacher effectiveness. *Journal of Teacher Education, 15*(1), 72-76.

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

*Hawk, P.P., & Schmidt, M.W. (1989). Teacher preparation: A comparison of traditional and alternative programs. *Journal of Teacher Education, 40*(5), 53-58.

Hedges, L.V. (1994). Fixed-effects models. In H.M. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis*. New York: Russell Sage Foundation.

Hutton, J.B., Lutz, F.W., & Williamson, J.L. (1990). Characteristics, attitudes and performance of alternative certification interns. *Education Research Quarterly,* 14(1), 38-48.

Ingersoll, R.M. (1999). The rroblem of underqualified teachers in American secondary schools. *Educational Researcher, 28*(2), 26-37.

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

*Knight, S.B., Owens, E.W., & Waxman, H. (1991). Comparing the classroom learning environments of traditionally and alternatively certified teachers. *Action in Teacher Education, 12*(4), 29-34.

*Laczko, I.I., & Berliner, D.C. (2000). *The effects of teacher certification on student achievement: An analysis of the Stanford Nine*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

*Laczko, I.I., & Berliner, D.C. (2001). *Does certification matter? An analysis of teacher certification on student achievement*. Paper presented at the annual meeting of the American Educational Research Association (AERA), Seattle, WA.

*Laczko-Kerr, I. (2002). *The effects of teacher certification on student achievement: An analysis of Stanford Nine achievement for students with emergency and standard certified teachers*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

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

Mandeville, G.K., & Liu, Q. (1997). The effect of teacher certification and task level on mathematics achievement. *Teaching and Teacher Education, 13*(4), 397-407.

*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.

National Association of State Directors of Teacher Education and Certification. (2002). *NASDTEC Manual*. Washington, DC: Author.

Perl, L.J. (1973). Family background, secondary school expenditures, and student ability. *Journal of Human Resources, 8*(2), 156-180.

*Pezzano, P.T. (1999). *A comparative study of traditional route and alternate route teachers*. Unpublished doctoral dissertation, Columbia University Teachers College.

Raudenbush, S.W. (1994). Random-effects models. In H.M. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis*. New York: Russell Sage Foundation.

Raymond, M., Fletcher, S.H., & Luque, J. (2001). *Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Unpublished report, CREDO, Stanford University.

*Sandlin, R.A.,Young, B.L., & Karge, B.D. (1993). Regularly and alternatively credentialed beginning teachers: Comparison and contrast of their development. *Action in Teacher Education, 14*(4), 16-23.

*Shim, C.-P. (1963). *A study of the cumulative effect of four teacher characteristics on the achievement of elementary school pupils.* Unpublished doctoral dissertation, University of Maryland.

*Shim, C.-P. (1965). A study of the cumulative effect of four teacher characteristics on the achievement of elementary school pupils. *Journal of Educational Research, 59*(1), 33-34.

Soar, R. S. (1983). Measures of quality in the classroom. *Journal of Classroom Interaction, 18*(2), 7-15.

Wu, M.-J., Becker, B.J., & Kennedy, M.M. (2002, April). Relationships of teacher qualifications and teaching quality: Conceptual issues and challenges. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

*Zirkle, J.E. (1998). *The effects of teacher certification on student achievement of music instruction goals and objectives*. Unpublished master's thesis.

Appendix

Table A1. *Coded Variables within each Comparison Type*

	Comparison	Trad vs Alt				Trad vs Emerg					
	Alt vs Emerg	Trad std vs Out				Trad prv vs Out					
		14	55	10	78	1	5	3	10	1	
	6										
Outcome											
1	Student Achievement Total	2	4	0	0	0	0	0	0	0	
	0										
2	Student Achievement Subtest	3	24	4	30	0	0	3	10	1	
	6										
3	Teacher Performance Total	2	2	1	2	0	0	0	0	0	
	0										
4	Teacher Performance Subtest	7	20	5	26	1	3	0	0	0	
	0										
5	Personality-like Measure	3	5	4	20	1	2	0	0	0	
	0										
Type of rater											
1	Inside observer	4	11	2	11	0	0	0	0	0	
	0										
3	Outside observer	4	13	4	37	1	5	0	0	0	
	0										
3	Student evaluation	1	3	0	0	0	0	0	0	0	
	0										
	No rater										
School level											
1	Elementary	5	32	5	29	1	5	1	2	0	
	0										
2	Secondary	1	2	3	32	0	0	2	8	1	
	6										
3	Mixed	7	21	3	17	0	0	0	0	0	
	0										
Subject											
1	Math	3	10	4	12	0	0	1	5	1	
	3										
2	Science	1	1	1	6	0	0	1	3	1	
	3										
3	Math or Science	1	5	0	0	0	0	0	0	0	
	0										

4	Reading	**3**	**5**	**3**	**6**	0	0	0	0	0	0
5	Language	**2**	**6**	**3**	**6**	0	0	0	0	0	0
6	Reading or Language	**0**	**0**	**0**	**0**	0	0	**1**	**2**	0	0
7	Arts and Science	**1**	**2**	**2**	**11**	0	0	0	0	0	0
8	Others	**10**	**26**	**4**	**37**	1	5	0	0	0	0
Experience control											
1	Non-controlled	**6**	**9**	**4**	**25**	1	5	**2**	**8**	1	6
2	Controlled	**10**	**46**	**6**	**53**	0	0	**1**	**2**	0	0
Publication type											
1	Journal	**7**	**25**	**3**	**30**	0	0	**2**	**8**	1	6
2	Dissertation	**4**	**26**	**3**	**17**	1	5	**1**	**2**	0	0
3	Report	**2**	**4**	**4**	**31**	0	0	0	0	0	0

*The left column under each comparison type shows the number of studies, the right column shows the number of effect sizes.

Table A2. *Analysis of Seven Comparison Groups*

	Q	df	p-value	Birge Ratio		
	Mean Effect		SE of the Mean			
Total	468.46**	152	<.0001	3.08	--	--
Between groups	108.80**	6	<.0001	18.13	--	--
Within groups	359.67**	146	<.0001	2.46	--	--
Trad prov vs. Alternate	151.03**		53	<.0001	2.85	-
0.05*	0.007					
Trad std vs. Alternate	0.01	1	.90	0.01	-0.03	
	0.136					
Trad prov vs. Emergency	131.3**	57	<.0001	2.30	0.13*	
	0.028					
Trad std vs. Emergency	48.83	19	.0002	2.57	0.08	
	0.055					
Trad prov vs. Out-of-field	7.40	5	.19	1.23	0.03	
	0.114					
Trad std vs. Out-of-field	13.68	7	.06	1.95	0.39*	
	0.053					
Alternate vs.Emergency	7.39	4	.12	1.85	-0.37*	
	0.171					

TM034990

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE
(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Does teacher certification imply quality?

Author(s): ~~B. BECKER~~ Qu, Yanxuan; B. BECKER

Corporate Source: Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

X Check here for Level 1 Release, permitting reproduction and dissemination in microfiche and other ERIC archival media (e.g. electronic) and paper copy.

or

_____ Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.

or

_____ Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

Sign Here, Please

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Position:

Printed Name: Yanxuan Qu Organization:

Address: 1640 K, Spartan Vlg. Telephone Number: (517)-353-6852

East Lansing, MI, 48823

Date:

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of this document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents which cannot be made available through EDRS).

Publisher/Distributor:

Address:

Price Per Copy: Quantity Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant a reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:

Address:

V.WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

You can send this form and your document to the ERIC Clearinghouse on Assessment and Evaluation. They will forward your materials to the appropriate ERIC Clearinghouse.

ERIC Acquisitions
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory (Bldg 075)
University of Maryland, College Park
College Park, MD 20742

(800) 464-3742
(301) 405-7449
eric_ae@ericae.net
http://ericae.net