An Examination of Teacher Quality Variables Associated with Passing State Content Tests

Pamela Esprivalo Harrell
Associate Professor in Secondary Education
University of North Texas, Denton
pam.harrell@unt.edu

Mary M. Harris
Professor of Teacher Education and Administration
University of North Texas, Denton
mary.harris@unt.edu

Jennifer K Jackson
Coordinator for Language Arts (grades 9-12)
Irving Independent School District
jjackson@irvingisd.net

Graduates of a post-baccalaureate secondary education program working on certification in five core subject fields served as a sample for transcript review to investigate how well the variables of number of hours completed for content major course work, the content major grade point average (GPA), and the age of content major course work predicted a passing score on the Texas state content certification test for teachers. The three predictor variables were statistically significant with GPA as the strongest single predictor associated with the Texas state content certification test.

Key Words: *Alternative Certification, content course work, content examinations*

During the 1980s, eight states offered alternative routes to teaching. In an effort to address an ever growing demand for teachers, particularly for high need assignments such as mathematics and science, increasing numbers of teacher candidates are presently prepared through these programs. Today, the number of

states offering alternative routes to teaching has expanded to include every state in the United States of America (Feistritzer, 2008).

The study reported here examines the relationship between the Texas Examinations of Educator Standards score and the following variables: (a) the number of upper division courses completed as part of a college major in the content field; (b) the grade point average on upper-level content courses; and (c) the number of months that elapsed between completion of the last upper-level content courses taken and the initial attempt at taking the state certification examination.

The authors are affiliated with a university-based alternative certification program that the state of Texas calls a "post-baccalaureate certification program." These students have obtained an undergraduate degree with a major or major equivalent in one of the subjects taught in the secondary schools. The program of study for these students enables them to complete 12 graduate semester credit hours of coursework in pedagogy through face to face or online learning environments, and enroll either in a semester of student teaching or a full year of supervised employment as the teacher of record in a secondary classroom. This post-baccalaureate program was designed for career changers, although many of the participants are fairly recent college graduates.

One difference between these teacher candidates and those in the teacher preparation programs of the 1980s is that their eligibility for certification will depend not only on completion of a degree or program of study leading to certification, but also on their performance on a state certification examination that includes subtests for content knowledge and pedagogy. Preparation for the pedagogy subtest is accomplished through the program delivered by the University of North Texas, but the content preparation of these teacher candidates was accomplished prior to program admission. This circumstance allows a unique opportunity to examine how well prepared these teachers were to become "highly qualified teachers" as defined by No Child Left Behind (NCLB), which requires both a content major (or equivalent of a major) and passing a content test. Three independent variables were examined: time lapse since courses were completed, the number of higher level content coursework or number of courses taken, and academic success of the teachers in these courses. The dependent variable was the participant's score on the content area licensure test

In summary, this study responds to the need to investigate content knowledge in the context of secondary certification programs. Specifically, the purpose of this study is to identify predictor variables for passing scores on the

Texas Examinations of Educator Standards content area examinations among candidates in a university-based post-baccalaureate certification program. Sections that follow review literature related to the role of certification testing within the framework of NCLB, describe the methods used in the study, the findings, and implications of the findings for teacher education programs and policy makers.

Review of Literature

Title II of *No Child Left Behind*, the sixth reauthorization of the *Elementary and Secondary Education Act of 1965*, offers the first federal definition of teacher knowledge that represents a major shift from the way teachers have traditionally been prepared and certified (Cohen-Vogel, 2005). NCLB significantly narrows the variables associated with teacher content knowledge by valuing a test score over other variables typically associated with teacher content knowledge, such as the numbers and types of college courses completed, the type of degree attained, the age of the coursework, grade point average, and verbal knowledge. A "highly qualified" teacher, as defined by NCLB, possesses a bachelor's degree and has passed a state test of content knowledge.

Teacher testing has been around in some form for about three decades. Russell (2005) marks the 1970s as the beginning of the trend toward formalized teacher testing in the United States. In an effort to promote high standards in the teaching profession and guarantee a minimum level of competency in the classroom, states began in the 1980s and 1990s to adopt strategies to ensure that only competent individuals entered the classroom. Teacher testing presented itself to policy makers as an efficient, cost-effective way to accomplish this goal (Russell, 2005). The Western Governors Association, a non-partisan coalition of governors serving 18 states and 3 U.S. flag Pacific Islands, was a pivotal player in the development of state content standards as the basis for teacher testing that could overcome variations in the state program approval standards for teacher education (Conley, 2003).

Although many states in the U.S. used some form of teacher testing, it was not until the enactment of NCLB that all 50 states and the District of Colombia required some sort of written test for initial teacher licensure. Presently, 43 states require assessments of basic skills, 43 require a test of subject matter knowledge in the certification field, and 35 require a test of pedagogical and professional

knowledge. Thirty states use all three forms of assessment, and 12 use two of these assessment methods (Council of Chief State School Officers, 2004).

This move toward teacher testing was an effort by policy makers to (a) improve the preparation of new teachers through the establishment of high state standards and accountability for initial teacher preparation and licensure, (b) reduce barriers to becoming a teacher among otherwise highly qualified individuals by retooling traditional teacher preparation programs and opening up alternative routes to teaching, and (c) to evaluate the content knowledge of teachers. A considerable body of research offers support for the introduction of teacher testing along with cautions about its limitations.

Support for Testing Teacher Academic Knowledge

Latham, Gitomer, and Ziomek (1999) champion the use of teacher tests as measures of teacher quality, citing teaching as an "academic enterprise" (p. 24). In addition, researchers (Kain & Singleton, 1996; Hanushek, 1971; Ferguson, 1991; Erenberg & Brewer, 1994; Latham, Gitomer, Ziomek, 1999; Laczko-Kerr & Berliner, 2003; Goldhaber, 2007; Clotfelter, Ladd, & Vigdor, 2006) have shown teachers' ability as measured by a variety of testing instruments to be positively correlated to their student's test scores.

Findings regarding teacher verbal ability were described by Hanushek (1971). He reported Stanford Achievement Test scores for 1,061 third graders in a California elementary school as significantly correlated with the *Quick Word Test Level 2* (New York: Harcourt, Brace and World, Inc., 1964) a verbal facility test given to the teachers.

Ferguson (1991) used data sets from grades one, three, five, seven, nine, and eleven in 900 schools across Texas to investigate the relationship between teacher scores on the Texas Examination of Current Administrators and Teachers (TECAT) and student achievement as measured using the Texas Educational Assessment of Minimum Skills (TEAMS) exam. Among his findings are that better literacy skills (i.e., higher TECATE scores) for teachers are statistically significantly related to higher student achievement scores. According to Ferguson, "Teacher's language skill as measured by the TECATE score is the most important school input for both math and reading (p. 475)."

Ehrenberg and Brewer (1994), using data from the *High School and Beyond* longitudinal study, investigated the influence of school and teacher characteristics on sophomore students' achievement test scores in mathematics,

vocabulary, and reading. Using institutional selectivity as a proxy for verbal ability or intelligence, they provided evidence that teacher association with more selective institutions was correlated with student gain scores on the tests.

Kain and Singleton (1996) used regression analysis to determine the effect of teacher test scores and student achievement in Texas schools, grades 3-7. The dependent variable used was the mean from the Texas Examination of Current Administrators and Teachers (TECAT) reading and writing scores. Scores derived from explanatory models using other teacher certification tests when TECAT scores were not available were also used. The variables examined included race/ethnicity, family income, and geographic location. Results indicate the teachers' mean TECAT reading score decreased as the numbers of black and Hispanic students increased. TECAT writing regressions indicated lower scores for teachers in schools with higher black and Hispanic student populations and lower numbers of high-income families. These finding support that "teacher ability, as measured by verbal and written proficiency scores, decreases as the campus percentage of black and Hispanic students increases; measured teacher ability increases with the campus percentage of high-income students" (p. 18).

Latham, Gitomer, and Ziomek (1999) investigated the academic preparedness of teachers using the scores of teacher candidates taking a Praxis, SAT, or ACT test from 1995-1997. Data included 34,000 Praxis I completers and 160,000 Praxis II completers who had also taken the SAT. The researchers found that candidates who pass the Praxis tests have academic skills comparable to or slightly better than the overall population taking the exam. An examination of those candidates seeking licenses in particular fields showed scores comparable to those in careers other than teaching. Passing rates by gender were similar, but gaps among ethnic groups were shown as "white candidates pass at a considerably higher rate than minority candidates" (p. 25).

Recently, other researchers have begun to examine the link between teacher quality variables and/or student achievement. Goldhaber (2007) examined the relationship between teacher tests and student learning gains as measured using a value-added system tied to the North Carolina Course of Study. His analysis included 24,237 grades 4-6 teachers in self-contained classrooms with either a NTE and/or Praxis test score and 722,166 students. Statistically significant effects were noted between teacher test scores and mathematics achievement. Goldhaber (2007) also explored the use of Praxis II mathematics tests as a screening device and found teachers who met the North Carolina standard to be more effective in teaching math (about 6% of a SD) compared to teachers who did not meet the standard. Another study conducted in North

Carolina found teacher licensure test scores had significant positive effects on students' 5th grade math scores (Clotfelter, Ladd, and Vigdor, 2006).

Cautions and Limitations

Along with studies that support the relationship between student achievement and teacher academic knowledge variables, some research evidence exists to support contentions that: (1) this relationship is not strong or is less important than other variables, (2) the teacher tests are inadequate measures of teacher knowledge, and/or (3) research is insufficient to justify policy based on the results of teacher tests. Main arguments of these positions are briefly reviewed here starting with the findings of several studies that challenge the relationship between teacher academic knowledge test results and student achievement.

Laczko-Kerr and Berliner (2003) explored the role of teacher quality in Arizona schools where one in six teachers was not certified. They found "the most consistent predictor of young students' achievement was the teacher's years of experience and not the teacher's subject matter competency" (p. 36).

Boyd, Grossman, Lankford, Loeb, and Wyckoff (2006) examined various New York City teaching pathways with regard to teacher test scores and student achievement in English Language Arts and mathematics. They found less than two percent of Teach for America (TFA) and Teaching Fellows failed the general knowledge certification exam on the initial attempt compared to 16% for College Recommended and Individual Evaluation teachers. However, in general, student achievement data for College Recommended teachers showed larger gains in student achievement for grades 4-8 during the initial years of teaching.

Angrist and Guryan (2004) examined the relationship of teacher grades, SAT scores, and Praxis scores. They concluded that grades are highly correlated with SAT scores, but not with scores on Praxis. This finding suggests Praxis is not screening for the same kind of academic characteristics as measured by grades and SAT scores. Additionally, no relationship between state exams and subject major was found, and testing was shown to have a statistically significant negative effect on the pass rate for Hispanic teachers.

Using data from the Schools and Staffing Survey, Angrist and Guryan (2007) investigated teacher quality for 160 teachers as related to the achievement of 3000 students. Their analysis employed various variables, including test scores (i.e., general and subject specific), teacher demographics, and teacher wage and quality measures (e.g., type degree awarded, average SAT score, type of

university attended, major in teaching subject). They concluded testing has no significant impact on teacher quality, although teacher testing requirements are associated with higher teacher wages, and testing is negatively associated with Hispanic representation in the new teacher pool.

Literature about teacher tests points up several types of limitations. In a review of state content tests for teachers then in use, including tests by both of the major national developers, Educational Testing Service and National Evaluation Systems, Mitchell and Barth (1999) concluded that most of the content assessed was at the high school level. They found the tests reviewed were psychometrically sound but seemed to reflect fear of litigation, resulting in a minimal competence approach. States that had adopted the same test selected different pass scores with variations as great as 14% in the number of correct answers required to pass. Moreover, the practice of reporting scores as "pass" or "fail" prevents study of the relationship between the strength of the passing score and other measures. Mitchell and Barth recommend that states choose the most rigorous licensure tests available, select credible pass scores, and require that candidates for teaching positions report their actual test scores to prospective employers and supply copies of their academic transcripts.

Other commentary has focused on the content validity of tests of teacher knowledge. Mitchell and Barth (1999) responded to the question, "What should teachers know and be able to do to teach their students to these (k-12) standards?" by stating that teachers require more advanced content knowledge than high school students in every domain of the content that is taught. The typical teacher test provides neither of these. The concept of "pedagogical content knowledge," introduced by Shulman (1986), suggests that teachers need, in addition to content knowledge, a grasp of the most likely misunderstandings of students, how to structure disciplinary content for teaching, and common ways of making disciplinary content comprehensible to students. Although Shulman recommends that pedagogical content knowledge be included in teacher tests (1986, p.10) the national and state content standards on which the tests are based vary in the extent to which they include this aspect. Russell (2005) summarizes research suggesting that another category of teacher knowledge not measured by tests, interpersonal and social skills, is more important than pedagogy and content knowledge test scores to success in the teaching and other helping professions. Darling-Hammond (2001) cites the tests of the National Board for Professional Teaching Standards as superior in validity to the more common pencil and paper tests.

Although a majority of states rely on teacher testing to ensure high standards, reliance on state academic subject testing to reflect teacher competency

is not generally supported by scholarship directed toward policy makers. Hanushek and Rivkin (2004) point out that, although content test scores are more often correlated with student outcomes that some other teacher characteristics, the focus and content of academic tests of teachers used by researchers vary too greatly for application in policy., In a critique of the Secretary of Education's first "Annual Report on Teacher Quality" following the passage of NCLB, Darling-Hammond and Youngs (2002) observe that although studies support that teacher subject matter knowledge is consequential, they do not justify "the Secretary's assertion that verbal ability and subject matter knowledge are more important for teacher effectiveness than knowing how to teach" (p. 19).

In a summary report for policy makers of the Education Commission of the States, Allen (2003) examined reports that met his criteria for rigor on the relationship between subject matter preparation of teachers and student test score results. This summary includes studies that find positive relationships between general academic knowledge and student success, such as Druva and Anderson (1983) and Monk (1994) and studies of the relationship of student performance to a teacher's having an undergraduate major in the discipline being taught, such as Ferguson and Womack (1993) and Goldhaber and Brewer (1997). Although not all studies showed positive relationships, Allen (2003) concluded that the research evidence is sufficient to support policies that require a measure of content knowledge for entry into teaching. This conclusion affirms the importance of teacher knowledge to the learning of children across socio-economic levels (Darling-Hammond, 2000). Also, measures of teacher knowledge seem especially apt for secondary teachers, for whom Allen suggests that policy makers require teachers to have an academic major in the discipline. Allen recommends that one purpose of a state teacher testing policy should be to diagnose content domains that may not have been acquired by the teacher candidates.

In fact, state licensure tests serve a gate-keeping rather than a diagnostic function. Because the requirement of a college major in the discipline leads one to imagine that the teacher tests are based on knowledge acquired through the completion of a bachelor's degree in the content field, strong alignment should exist between university coursework and licensure tests for secondary teacher candidates. This association advances considerably the importance of the research question:

> **Research Question.** What is the relationship of the Texas Examinations of Educator Standards score to the number of upper-level content area courses; (2) upper-level content area grade point average; and

(3) the number of years between the last upper-level content area course completed and the month the teacher candidate initially attempted the test?

Method

Sample

Tables 1-4 present demographic information about the 287 students enrolled in the Online Post-Baccalaureate Secondary Program between September 1, 2004, and January 1, 2008. About eighty-one percent of the participants were white. Sixty-one percent were female (Table 1 and Table 2)

Table 1
Ethnicity of Participants

N=287	Number	%
African American	10	3.5
Asian	3	1.1
Hispanic	12	4.2
Native American	7	2.4
White	235	81.9
Not Specified	20	7.0

Percentages may not equal 100% due to rounding error.

Table 2
Gender of Participants

N=287	Number	%
Female	176	61.3
Male	111	38.7

Candidates earned degrees from many accredited colleges and universities which ranged from open admission to highly selective colleges and universities, according to *U.S. News and World Report*. Seventy percent of the individuals had earned a bachelor's degree from a college or university in Texas, including 31.71% from the University of North Texas; 39.02% of them had earned a bachelor's degree outside Texas (Table 3).

Table 3
Location of Baccalaureate Degree of Participants

N=287	Number	%
Baccalaureate degree from other states	112	39.02
Baccalaureate degree from Texas institutions	175	69.97
Baccalaureate degree from the University of North Texas	91	31.71

Percentages may not equal 100% due to rounding error.

The candidates were distributed among the secondary content fields, with 34.2% having earned degrees in English, 23.3% in history, 15.7% in life science, 12.9% in mathematics, and 13.9% in social studies (Table 4).

Table 4
Certification Area of Participants

N=287	Number	%
English Language Arts and Reading	98	34.2
History	67	23.3
Life Science	45	15.7
Mathematics	37	12.9
Social Studies	40	13.9

As program completers, each of these students had met the requirements for admission to the University of North Texas graduate certification program and completed 12 hours of pedagogically focused coursework as well as completion of student teaching or a year-long teaching practicum. None of these requirements were specifically directed to improving candidate content knowledge or preparing candidates for the content sub-test of the Texas Examinations for Educator Standards (TExES).

Instrumentation

Academic transcripts of these students were used to determine the level and number of upper-level content area courses completed, the numbers of upper-level credits earned, and the year and month the last upper-level content area course was completed. Also noted were the name of the university from which each candidate graduated and the titles, course numbers, and grades earned in each of the upper-level content courses, This academic transcript information is routinely collected for evaluation purposes by the administrator for the University of North Texas Post-Baccalaureate Secondary Program. Reliability was established using a pool of three experts who coded content courses by type.

Agreement was established by two experts being in agreement. A third expert was used as a tie-breaker.

The English Language Arts and Reading, History, Life Science, Mathematics, and Social Studies content tests of the TExES were taken by the individuals seeking initial teacher certification through the Post-Baccalaureate Secondary Program. The score used for content assessment was the score earned by each candidate on the first test attempt.

Data Analysis

A linear regression analysis was performed using the candidate's scaled score on the Texas Examinations of Educator Standards as the dependent variable. Scores of 240 or above are considered passing. The statistical package SPSS ® version 15 was used for all statistical analyses. The regression (B) coefficients and the correlation coefficients (R), (R^2), and adjusted (R^2) are presented. Effect sizes of each predictor variable are discussed. Prior to running the analysis, histograms and scatter plots were used to ensure that the assumptions of multiple regression had been met.

Data were collected from the 287 students who were enrolled in the Secondary Online Teacher Certification Program during the period of September 2004 through January 1, 2008 and who completed the Texas teacher content test (TExES). Descriptive data and multiple linear regression statistical analyses provided findings about the subject area knowledge of these individuals and the significance of the predictors examined. Descriptive data included the means, modes, and standard deviations of the variables included in this study. Linear regression analysis was used to identify predictor variables associated with passing five core subjects for the grades 8-12 Texas Examinations for Educator Standards for content in English Language Arts (ELA), History, Life Science, Mathematics, and Social Studies.

The primary dependent variable was the score on the Texas Examinations of Educator Standards. This variable was used to examine the relationship between the TExES score and the three predictor variables: (a) upper-level content area grade point average, (b) number of upper-level content area courses taken, and (c) the number of months that elapsed between completion of the last upper-level content course completed and the initial attempt at taking the certification examination.

Results

Descriptive statistics

Tables 5, 6, and 7 show descriptive statistics for the months of time elapsed between the last upper-level content-area course completed and the three predictor variables (i.e., the initial attempt on the TExES, the upper-level content area grade point average, and the number of upper-level content-area courses taken).

Mixed results have been reported regarding the representation of career changers in alternative certification programs. It is interesting to note that, with the exception of the mathematics candidates in this study, most of the candidates were recent graduates. As shown in Table 5, the median scores for ELA, history, life science, and social studies candidates suggest that at least half of the candidates in this program graduated within the last two years and mean scores show the average candidate graduated within the last five years. The most frequent number of months between the last upper-level content-area course completed and the initial attempt on the TExES was less than seven months for this group of candidates. However, half of the mathematics candidates graduated 7.5 or more years prior to entering the program and the mean for this group was 136.22 months (11.3 years). In general, the mathematics candidates represent a career changer group.

Table 5
Descriptive Statistics for Months of Time Elapsed between Upper-level, Content-area Coursework Completed and the Initial TExES Attempt

	N	Min	Max	Median	Mean	SD	Mode
ELA	98	.00	348	15.50	40.91	62.99	12
History	67	.00	369	17.00	51.03	17.000	17
Life Science	45	2.00	348	17.00	42.60	69.432	4
Mathematics	37	2.00	401	91.00	136.22	124.604	11
Social Studies	40	.00	202	18.50	35.45	43.173	8
All Groups	287	.00	401	19.00	55.06	83.277	5

In addition to the age of coursework, GPA and the number of content courses completed have traditionally been used by teacher preparation programs as measures of teacher quality. In general, the median and mean GPAs suggest that candidates in this program were well-prepared, having median and mean GPAs above 3.0, where 4.0 = A (Table 6). Almost 80% of candidates had upper-level content GPAs above 3.0. Furthermore, the medians and means reported

indicate that most of the candidates completed at least six upper-level courses, with life science candidates completing the most upper-level coursework overall (Table 7). A possible explanation for the relatively high number of upper-level life science courses is the practice of many biology programs supporting the preparation of premed students with substantial coursework in biology through their majors. Many ELA candidates held a master's degree, which accounts for higher numbers of content coursework for this group. Two outliers present two upper-level courses, which indicates an error in program admission as candidates must have at least 12 hours of upper-level coursework for program admission.

Table 6
Descriptive Statistics for Upper-level Content Grade Point Average

	N	Min	Max	Median	Mean	SD	Mode
ELA	98	2.50	4.00	3.50	3.50	.374	4
History	67	2.00	4.00	3.34	3.31	.462	3
Life Science	45	2.00	4.00	3.03	3.10	.506	3.5
Mathematics	37	1.00	4.00	3.14	3.03	.727	4
Social Studies	40	2.29	4.00	3.49	3.40	.415	4
All groups	287	1.00	4.00	3.38	3.38	.507	4

Table 7
Descriptive Statistics for Total Number of Upper-level Content-area Courses Completed

	N	Min	Max	Median	Mean	SD	Mode
ELA	98	2.00	28.00	9.00	9.70	4.616	6
History	67	2.00	20.00	7.00	7.36	3.450	8
Life Science	45	4.00	24.00	11.00	11.67	4.266	12
Mathematics	37	2.00	15.00	7.00	6.59	3.122	4
Social Studies	40	4.00	22.00	10.00	10.30	4.421	7
All Groups	287	2.00	28.00	8.00	9.05	4.373	8

Although the descriptive statistics suggest that program participants are knowledgeable about the content associated with their respective certification fields, the number of individuals who passed or failed their respective Texas Examinations of Educator Standards content tests might suggest the knowledge acquired through university coursework is not in sync with the state content test (Table 8). Results indicate 54 of the 287 students (18.8%) failed their respective Texas Examinations of Educator Standards content test on the first attempt. The ELA group had the highest passing rate with 93 individuals passing the test on the first attempt. The Mathematics test group had the lowest passing rate with 22 out of 37 students (59.5%) passing the test on the first attempt. The Life Science test group had a 68.9% pass rate on the first test attempt.

Table 8 presents descriptive data on the Texas Examinations of Educator Standards scaled scores. Results indicate that the ELA group had the highest

Table 8
Descriptive Statistics for TExES Scaled Test Score by Certification Field Test

	ELA	History	Life Science	Mathematics	Social Studies	Total
N	98	67	45	37	40	287
Min	0	213	193	211	219	0
Max	295	288	291	292	282	295
Mean	268	242	244	238	230	268
Median	259.65	255.67	248.98	253.84	254.48	255.58
SD	267.00	257.00	251.00	254.00	255.00	261.00
#Fail	39.700	18.813	25.681	23.640	16.452	28.949
%Fail	5	13	14	15	7	54
#Pass	5	19.4	31.1	40.5	17.5	18.8
%Pass	93	54	31	22	33	233
Total	95	80.6	68.9	59.5	82.5	81.2

mean TExES score, 259.65; the highest median TExES score, 267; and the most variability compared to the other groups (SD = 39.700). The Life Science group had the lowest mean TExES score, 248.98; the lowest median TExES score, 251; and the second most variability (SD = 25.681).

Regression Analysis

English Language Arts. The sample size for English Language Arts (ELA) candidates is sufficient to conduct an independent linear regression

analysis involving three independent variables. The analysis of variance results for the three-predictor model was statistically significant: $F(3, 94) = 4.320$, $p < .007$. The summary of regression for the English Language Arts and Reading Test Group shows the three-predictor model was statistically significant and identifies GPA as a significant predictor of the TExES test score (Table 9 and Table 10).

Table 9
Summary of Regression Model for ELA using Three Predictor Variables

Test Group	R	R^2	Adjusted R^2	Sig	Std. Error of the Estimate
English Language Arts and Reading (N = 98)	.348	.121	.093	.007	37.807

The regression equation is:
TExES score = 156.554 + 28.970 (GPA) + .604 (# Courses) -.098 (# months elapsed).

The overall effect size for the ELA group was $R^2 = 0.121$. That is, all of the independent variables explained 12% of the variance.

Table 10
Coefficients for Three-predictor Linear Regression Model for the English Language Arts and Reading Test Group

Model	Unstandardized Coefficient		Standardized Coefficients			
	B	Std. Error	Beta	t	Sig.	
Constant	156.554	37.244		4.204	.000	
GPA	28.970	10.579	.273	2.738	.007	*
Courses	.604	.853	.070	.708	.481	
Months Elapsed	-.098	.063	-.156	-1.564	.121	

* Statistically significant (P<.05)

A multiple correlation analysis for each independent variable (Table 11) reveals that GPA and test score are significantly related ($r = .311$; $p < .01$). The number of months of time elapsed between the last upper-level content-area course taken and the initial attempt on the ELA TExES and the TExES score are negatively related, but not statistically significant ($r = -195$), and the number of upper-level content-area ELA courses taken and the TExES score are not statistically significant ($r = .084$).

Table 11
Correlations between Scores on the English Language Arts and Reading TExES Test and Three Predictor Variables

Variable	TExES Score	# Content Courses	Upper-level GPA	Months Elapsed
TExES Score	1			
# Content Courses	.084	1		
Upper-level GPA	.311*	.135	1	
Months Elapsed	-.195	.148	-.180	1

* Correlation is significant at the 0.01 level (2-tailed)

All Core Content Fields. An examination of the data for all core content fields show the analysis of variance results for the three-predictor model was statistically significant, $F(3, 283) = 9.034$, $p < .000$. The summary of regression for All Fields shows the three-predictor model was statistically significant and identifies GPA as a significant predictor for TExES test score (Table 12 and Table 13).

Table 12
Summary of Regression Model for All Fields for Three Predictor Variables

Test Group	R	R^2	Adjusted R^2	Sig	Std. Error of the Estimate
All Fields (N = 287)	.296	.087	.078	.000	27.801

The regression equation is:
TExES score = 198.003 + 15.897 (GPA) + .464 (# Courses) .012 (# months elapsed).

The overall effect size for All Test Groups was $R^2 = 0.087$. That is, all of the independent variables explained 8.7% of the variance.

A multiple correlation analysis for each independent variable (Table 14) reveals that GPA and test score are significantly related ($r = .286$; $p < .01$) as are the number of content courses taken and GPA ($r = .174$; $p < .01$). There is a negative correlation between the upper-level GPA and the months elapsed since completion of the content degree ($r = -.120$; $p < .043$). The number of months of

time elapsed between the last upper-level content-area course taken and the initial attempt on the TExES are negatively related, but not statistically significant (r = -.008), and the number of upper-level content-area courses taken and the TExES score are not statistically significantly related (r = .115; p = .052).

Table 13
Coefficients for Three-Predictor Linear Regression Model - All Core Field Test Groups

Model	Unstandardized Coefficient		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	198.003	11.293		17.534	.000
GPA	15.897	3.313	.278	4.798	.000
Courses	.464	.384	.070	1.209	.228
Months Elapsed	.012	.020	.034	.594	.553

Table 14
Correlations between Scores for All Fields and Three Predictor Variables

Variable	TExES Score	Number Content Courses	Upper-level GPA	Months Elapsed
TExES Score	1			
# Content Courses	.115	1		
Upper-level GPA	.286*	.174*	1	
Months Elapsed	-.008	-.115	-.120*	1

** Correlation is significant at the 0.05 level (2-tailed)
*Correlation is significant at the 0.01 level (2-tailed)

Conclusions and Implications for Practice

This study's findings suggest statistically significant relationships between the number of upper division (higher degree of difficulty) content courses completed, the upper division content GPA, and the number of months elapsed between content completion and the first attempt on the content area licensure test. Although the effect size was small, we conclude that these variables may predict content licensure test scores. Of the three variables examined, only GPA was a statistically significant predictor for success on the TExES examination.

The sample numbers for core field subgroups permitted an examination for English Language Arts candidates and for all fields combined. However, individual field samples were insufficient to analyze fields other than English Language Arts, a limitation of the study.

Ninety-five percent of ELA candidates passed the content examination, and this group tended to have completed a comparatively large number of courses (Mean = 9.70) with the highest mean GPA of the content subgroups. Completing an even higher number of upper division courses (Mean = 11.833), Life Science candidates passed the TExES at a rate of 66.6%, The Mathematics pass rate was even lower, 57.6%, with this group showing the greatest amount of time elapsed between upper division course completion and first attempt on the TExES. Science and mathematics are shortage fields in Texas. Policy makers recently changed minimum Texas high school graduation requirements to include four years of mathematics and science. Possibly, the test acts as a gatekeeper to certification for individuals with degrees in mathematics or science who might otherwise become certified.

All of the TExES tests were recently developed with a multimillion dollar budget and were designed specifically for Texas teachers. The TExES tests replaced the ExCET test system in 2004. The TExES tests are purported to align with the Texas Essential Knowledge and Skills (TEKS), the state curriculum for K-12 students, which were adopted in 2002-2003. It is assumed that study of the state curriculum defined by the TEKS will prepare a student for college from which the student will eventually graduate with an academic major. In the case of secondary teachers, the academic major is required for teacher certification and should relate to the certification field for which the individual receives a teaching license. Since NCLB mandates a bachelor's degree for alternative certification candidates, it is assumed that the content courses taken during the degree will prepare the candidate to pass the state content exam. Given the assumption of NCLB, it is surprising that content coursework taken as part of an academic degree is not sufficient to pass a teacher content examination. Although English Language Arts teachers took many upper-level courses accompanied by high GPA scores , the presence of coursework and GPA in other content fields did not seem to be necessary and sufficient to pass the TExES exam. For the content field of mathematics and science, the TExES functioned as a gatekeeper to certification in spite of strong content preparation and GPA; however, a larger sample population would be required to strongly support this inference.

Mathematics candidates were least likely to pass the state contents on the first attempt. Mathematics candidates differed from those in the other fields

studied by representing more genuine career changers, individuals who had completed their content major years ago. Also, these candidates had completed the fewest credits of advanced coursework in their majors and had the second lowest mean GPAs. A wide variety of conclusions could be drawn from these factors. Perhaps, the mathematics candidates did poorly on the content exams because of the age of their coursework. Or, perhaps, their college majors, taken before the days of standards-based accountability in education, were less aligned with the K-12 curriculum than the majors of more recent college students. Perhaps these college mathematics majors took less upper division coursework because of poorer K-12 preparation for college compared to students in other disciplines. Or perhaps these candidates who eventually entered teaching were less able than their mathematics major peers, contributing to their taking fewer advanced courses and/or to their low grades. Or, perhaps, college mathematics faculty grade harder than faculty in other disciplines. From the evidence available, it is not possible to tell which of these explanations are likely to inform policy solutions.

Second least likely to perform successfully on the state content test were the Life Science candidates in spite of the fact that this group tended to present the largest number of advanced courses in the major. Unlike the mathematics candidates, the life science group tended to be comprised of recent college graduates. Among the five groups studied, this group presented the lowest mean GPA (3.10). For the Life Science group, the arguments suggested for the Mathematics group about age of coursework's accounting for lack of alignment with education standards and lack of upper division coursework's being related to the lesser ability of these particular majors appear not to be relevant. Arguments include that the college major is not aligned to the standards, that these candidates were not star performers, and/or that faculty in this discipline, which may cater to medical students, grade hard. It should be noted, however, that the Post-baccalaureate Secondary Education program requires a minimum 2.8 GPA for admission, so weaker candidates, as measured by GPA, are automatically excluded from this program.

Policy makers may need to examine the assumption that teacher testing should represent an exclusive method for assessing teacher content knowledge. This study suggests large numbers of candidates holding degrees from major universities, with accompanying high GPAs, and significant and/or recent content coursework, experience delays in completing certification requirements due to failing a state content exam. With the exception of the mathematics candidates and a few outliers, the descriptive statistics for time elapsed since the last upper-level content-area course was taken, upper-level GPA, and number of content-

area courses completed suggest that the program candidates demonstrated strong and recent content knowledge. Other approaches suggested by the literature include that teacher testing serve a diagnostic purpose (Allen, 2003) and that teacher test scores be considered with other measures, such as the comprehensiveness of the content major (Mitchell & Barth, 1999) in teacher certification.

The fact that 18.8% of post-baccalaureate secondary teacher candidates in core subject areas failed state content examinations, with the highest failure rates in high need fields, such as mathematics and science, deserves further examination. When high GPA and recent and ample content coursework lead to failure rates of 31.1% for life science and 40.5% for mathematics teacher candidates, one must question using a state exam as the sole gatekeeper for assessing teacher content knowledge. Although Texas has long used teacher testing as a form of accountability, it appears that changing the gatekeeper for teacher content knowledge from coursework completed at colleges and universities to a state teacher content examination presents a significant barrier to achieving teacher certification. Tests do provide a simple form of accountability that can be viewed as both efficient and cost effective. However, it may be that using either a content degree or a teacher test score might present an efficient and cost efficient method to assess teacher content knowledge with the bonus of moving teachers through the pipeline at an accelerated rate because they do not need to schedule a teacher content exam one or more times.

References

Allen, M. (2003). *Eight questions on teacher preparation: What does the research say*? Denver, CO: Education Commission of the States.

Angrist. J. D. & Guryan, J. (2007). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review, 27*, 483-503.

Angrist, J. D. & Guryan, J. (2004). Teacher testing. Teacher education, and teacher characteristics. *American Economic Review, 94*(2), 241-46.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006, April). How changes in entry requirements alter the teacher workforce and affect student achievement. Albany NY: Teacher Policy Research. Retrieved on August 26, 2009 from http://www.teacherpolicyresearch.org/portals/1/pdfs/how_changes_in_entry_requirements_alter_the_teacher_workforce.pdf

Clotfelter, C. T., Ladd, H. F. & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources 41*(4): 778–820.

Cohen-Vogel, L. (2005). Federal role in teacher quality: "Redefinition" of policy alignment? *Educational Policy, 19*(1), 18-43.

Conley, D. (2003). *Who governs our schools: Changing roles and responsibilities*. New York: Teachers College, Columbia University.

Council of Chief State School Officers (2004). *Key state education policies on PK-12 education: 2004.* Washington, DC: Council of Chief State School Officers.

Darling-Hammond, L. (2001). Teacher testing and the improvement of practice. *Teaching Education 12* (1), 11-34.

Darling-Hammond, L. (2000). How teacher education matters. *Journal of Teacher Education, 51*(3), 166-173.

Darling-Hammond, L. & Youngs, P. (2002). Defining "highly qualified teachers": What does scientifically-based research actually tell us? *Educational Researcher, 31* (9), 13-25.

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

Ehrenerg, R. C. and Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review 13*(1) 1-17.

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

Ferguson, R. F. & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education. 44*(1), 55.63.

Feistritzer, C. E. (2008). Alternative teacher certification: A state-by-state analysis 2008. Washington, DC: National Center for Education Information.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *The Journal of Human Resources 42*(4), 765-794.

Goldhaber, D. D. & Brewer, D. J. (1997). Teacher licensing and student achievement. In M. Kanstoroom, & C. E. Finn (Eds.), *Better Teachers Better Schools* (pp. 83-103). Washington, DC: Thomas B. Fordham.

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

Hanushek, E. & Rivkin, S. (2004). *How to improve the supply of high-quality teachers.* Brookings Papers on Educational Policy, 7-25.
Kain, J. F. & Singleton, K. (1996). Equality of educational opportunity revisited. *New England Economic Review*, May/June 87-111.
Laczko-Kerr, I. and Berliner, D. C. (2003). In harm's way: How undercertified teachers hurt their students. *Educational Leadership, 60*(8), 34-39.
Latham, A. S., Gitomer D., & Ziomek, R. (1999). What the tests tell us about new teachers. *Educational Leadership, 56*(8), 23-26.
Mitchell, R., & Barth, P. (1999). How teacher licensing tests fall short. *Thinking K-16*. Washington, DC: The Education Trust. Retrieved August 12, 2009, from http://www.edtrust.org
Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13(2),* 125-145.
Quick Word Test Level 2 (1964). New York: Harcourt, Brace and World, Inc.
Russell, T. (2005). Preparing for teacher testing in Ontario. Retrieved August 12, 2009, from http://educ.queensu.ca/~russellt/w2000/ttesting.htm
Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 5*(2), 4-14.