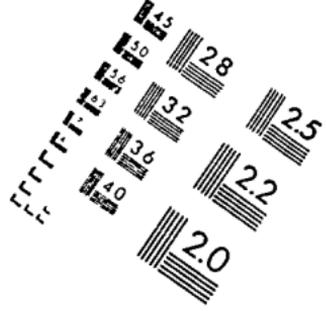
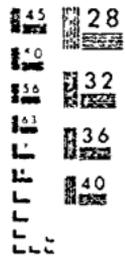




ERIC  
Full Text Provided by ERIC



ABCDEFGHIJKLMNQRSTUUVWXYZ  
abcdefghijklmnopqrstuvwxyz1234567890

BCDEFGHIJKLMNOPQRSTUVWXYZ  
abcdefghijklmnopqrstuvwxyz  
1234567890

DOCUMENT RESUME

ED 270 389

SP 027 007

**AUTHOR** Garcia, Peter A.  
**TITLE** A Study on Teacher Competency Testing and Test Validity with Implications for Minorities and the Results and Implications of the Use of the Pre-Professional Skills Test (PPST) as a Screening Device for Entrance into Teacher Education Programs in Texas.

**SPONS AGENCY** National Inst. of Education (ED), Washington, DC.  
**PUB DATE** Aug 85  
**GRANT** NIE-G-85-0004  
**NOTE** 144p.  
**PUB TYPE** Reports - Research/Technical (143)

**EDRS PRICE** MF01/PC06 Plus Postage.  
**DESCRIPTORS** Academic Standards; Educational Legislation; \*Minimum Competency Testing; \*Minority Groups; Teacher Certification; \*Teacher Evaluation; \*Test Reliability; \*Test Validity

**IDENTIFIERS** Texas

**ABSTRACT**

The educational reform movement has brought a mandate to test prospective teachers from training entry through exit and certification. Concentrating upon recent literature, this report explores current trends and policies in teacher testing. A review of what tests measure and how tests are being used is included. The negative impact of testing on the minority teaching populations is stressed. Underrepresented in the current teaching force, ethnic minorities are projected to become majority student populations in states like Texas and California. Causes of failure of ethnic minorities, remediation efforts in some states, and legal decisions which have played a major role in the reform movement and in testing are reviewed. Solutions being applied to the testing reform movement are discussed with futuristic planning which could bring about better cooperative efforts in a national climate of great diversity in the 50 states. The results of a national survey on the use of teacher tests and their impact on ethnic minorities are included. The final part of the study is a review and critique of the validation process done in Texas to support the use of the Pre-Professional Skills Test as an entrance examination for teacher education programs. Conclusions are drawn on legal and ethical issues in testing and the fairness of current tests in regard to minority teachers. A 17-page list of references, charts and graphs providing information on the current status of testing in the United States, and the survey instrument are appended. (Author/JD)

\*\*\*\*\*  
 \* Reproductions supplied by EDRS are the best that can be made \*  
 \* from the original document. \*  
 \*\*\*\*\*

**A Study on Teacher Competency Testing and  
Test Validity with Implications  
for Minorities  
and  
The Results and Implications of the Use of the  
Pre-Professional Skills Test (PPST)  
as a Screening Device for Entrance  
into Teacher Education Programs in Texas**

**U.S. DEPARTMENT OF EDUCATION  
NATIONAL INSTITUTE OF EDUCATION  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)**

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

---

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

by

**Peter A. Garcia  
Pan American University  
Edinburg, TX 78539-2999**

A Study on Teacher Competency Testing and  
Test Validity with Implications  
for Minorities  
and  
The Results and Implications of the Use of the  
Pre-Professional Skills Test (PPST)  
as a Screening Device for Entrance  
into Teacher Education Programs in Texas

Final Report

NIE Grant No. NIE-G-85-0004

Submitted to:

The National Institute of Education

by

Peter A. Garcia  
Pan American University  
Edinburg, Texas 78539

August 1985

## Table of Contents

	Page
Abstract. . . . .	1
Part I: Review of Literature. . . . .	2
The Present . . . . .	2
Teacher Competence . . . . .	8
What Tests Measure . . . . .	11
Major Tests in Use . . . . .	16
The National Teacher Examination. . . . .	17
The Pre-Professional Skills Test. . . . .	18
The California Basic Educational Skills Test. . . . .	19
Table 1--Percent of Examinees Passing the CBEST in the First Year of Administration. . . . .	21
The Scholastic Aptitude Test. . . . .	21
Use of Tests . . . . .	23
Teacher Test Use. . . . .	23
Misuse of Tests . . . . .	26
Guidelines for Proper Test Use. . . . .	27
The Impact of Testing on Minorities. . . . .	33
Impact on Teacher Education Programs. . . . .	35
Impact of Multi-lingual Populations . . . . .	36
Impact on Teaching Populations. . . . .	37
Cause for Failure. . . . .	37
Hispanic Population Data . . . . .	39
Legal Aspects. . . . .	41
Solutions. . . . .	45
The Future . . . . .	48
Part 2: National Survey on Testing. . . . .	51
Background. . . . .	51
Targeted Areas. . . . .	51
Perceptions of Test Content . . . . .	52
Table 2--States' Responses to "What Does the Examination Test?". . . . .	53
Test Use as Reported by States. . . . .	54
Table 3--States Response to Content/Use Categories for Each Test Listed for Target Areas of Competency. . . . .	55

Table of Contents (continued)

	Page
Impact on Minority Groups . . . . .	60
Table 4--Pass/Failure Rates by Ethnic Group. . . . .	60
Table 5--A Comparison of White vs. Ethnic Group Pass Rates. . . . .	61
Questions of Multiple Criteria. . . . .	62
 Part 3: An Analysis of the PPST as used in Texas. . . . .	 63
Introduction. . . . .	63
Determining PPST Suitability. . . . .	64
Adequacy-of-Preparation. . . . .	64
Content Validity . . . . .	68
Content Validity in Terms of Job Relevance and Program Relevance . . . . .	69
Performance Standards. . . . .	72
Determining PPST Impact . . . . .	74
Table 6--Ethnicity of Texas Full-Time Personnel for 1983-84 . . . . .	75
Table 7--Ethnicity of Texas Public School Children from 1981-83 . . . . .	75
PPST Results to Date . . . . .	76
Table 8--Cumulative Passing Rate for Students Who First Took PPST in March 1984. . . . .	77
Table 9--PPST Longitudinal Data from March 1984 through March 1985. . . . .	79
Testing of Current Educators With the TECAT . . . . .	79
 Conclusions . . . . .	 82
 Recommendations . . . . .	 85
 References. . . . .	 87
 Appendixes	
Appendix A. . . . .	104
Appendix B. . . . .	105
Appendix C. . . . .	106
Appendix D. . . . .	107
Appendix E. . . . .	108
Appendix F. . . . .	109
Appendix G. . . . .	110

Table of Contents (continued)

	Page
Appendix H. . . . .	.111
Appendix I. . . . .	.112
Appendix J. . . . .	.113
Appendix K. . . . .	.122
Appendix L. . . . .	.123
Appendix M. . . . .	.124
Appendix N. . . . .	.125
Appendix O. . . . .	.126
Appendix P. . . . .	.127
Appendix Q. . . . .	.128
Appendix R. . . . .	.129
Author Notes. . . . .	.130

## Abstract

The educational reform movement sweeping the country has brought with it a mandate to test prospective teachers from training program entry through exit and certification. Testing is essentially an accepted and inexpensive method of "measuring" knowledge or characteristics associated with competent teachers. Concentrating upon most recent literature, this report explores current trends and policies. A review of what tests measure and how tests are being used is included. The negative impact of testing on minority teaching populations is stressed. Underrepresented in the current teaching force, ethnic minorities are projected to become majority student populations in states like Texas and California.

Causes of failure of ethnic minorities are reviewed as are remediation efforts in some states. Legal decisions which have played a major role in the reform movement and in testing are reviewed. Solutions being applied to the testing reform movement are discussed with futuristic planning which could bring about better cooperative efforts in a national climate of great diversity in the fifty states.

The results of a national survey on the use of teacher tests and their impact on ethnic minorities are included. The final part of the study is a review and critique of the validation process done in Texas to support the use of the PPST as an entrance examination for teacher education programs.

## Part 1

### Review of Literature

#### The Present

There is a genuine concern on the part of the American people about the preparation and effectiveness of teachers in American schools. In response to this concern, the current teacher reform movement continues to escalate at a rate that challenges the extent to which a complete monitoring system can be devised. The National concern about the condition of public education has generated steadfast support for accountability in the education establishment. Sanders, Benton, Kaagan, Simons, and Teague (1984) acknowledge the problem that some teachers have not been drawn from the segment of American students with the greatest academic ability. They also cite arguments relative to standard measurement processes as being reasonable indicators of teaching ability and a compelling need for incentives to attract high quality teachers. Benderson (1982) states that the tests taken by teachers in Dallas, Houston, and Mobile, Alabama indicate that in some cases teachers were hardly more skilled in the basics than are their students.

State and national policy makers have undertaken the task of mandating changes in public education intended to improve the quality of teacher education in the country since a series of education commissions has indicated that the manner in which teachers are trained is inadequate. Hall (1985) and Hoffman and Defino (1985) indicate that there is little research evidence available for or against mandates.

Widespread reports that education majors are not as academically able today as compared to the past led to regional and national reports citing the shortcomings of educators. The charges made by Sikula and Roth (1984) of dropping SAT scores and lack of academic rigor in education draw attention to admission criteria of teacher education programs. Partially in response to the April 1983 release of the report A Nation at Risk: An Imperative for Educational Reform (1983), the American Association of State Colleges and Universities (AASCU) conducted a survey of state colleges and universities to learn how they worked with elementary and secondary schools in solving problems. Of the 64% respondents, requirements for entry into teacher education programs varied as indicated by the following findings:

A mean grade point average of 2.4 . . . 13 percent required a preprofessional exam; 34 percent required an orientation course; 73 percent required an English competency exam. Recommendations, interviews, and a physical examination were also necessary for admittance into several programs. ("Association Briefs," 1985, p. 19)

A survey conducted by Garcia in 1985 (which comprises part two of this report) supports the results of the AASCU (Sikula, 1984) study by showing that the great majority of respondents continue to use multiple criteria for admission into teacher education. (Only one indicated the use of a single cut-off score on a test as the sole criterion.) However, all 16 states which reported data on failure rates of ethnic minorities confirmed, through a telephone survey, the present use of established cut-off scores on state mandated tests which prevent entry or continuance within teacher education programs or which may be used to withhold certification. Multiple criteria notwithstanding, single cut-off scores are very powerful. Some states reported a limited number of attempts afforded testees to pass the examination with an acceptable cut-off score while others provided unlimited opportunities to achieve the cut-off scores.

The survey done by AACSU showed that 56% of the institutions require teacher education graduates to pass a competency exam. Sandefur (1985) reported that 76 percent of the states are now requiring some form of competency assessment (see Appendixes A and B for complete data).

At its July 3, 1985 national convention, the 7,500 delegates of the 1.7 million member National Education Association (NEA) reiterated opposition to competency tests for practicing teachers. However, the delegates endorsed a resolution supporting state certification exams, graduating GPAs at the B- level, and an internship program for new teachers. A Gallup Poll commissioned by NEA released July 1, 1985, continues to verify strong public support for competency tests for all teachers. The American Federation of Teachers (AFT) and NEA oppose competency tests for experienced teachers (Ferraro, 1985).

Due to the widespread belief that teacher education majors are not as well prepared as in the past, the Nation at Risk report recommends the following changes to the nation's institutions of higher learning (cited in Sikula and Roth, 1984):

- . Schools, colleges, and universities adopt more rigorous and measurable standards and higher

- expectations for academic performance and student conduct;
- . Four-year colleges and universities raise their requirements for admission; and
  - . Persons preparing to teach be required to meet high educational standards, to demonstrate an aptitude for teaching, and to demonstrate competence in an academic discipline. (p. 9)

The commission made popular several ideas believed by educators for some time. For example, the report stated that:

- . Not enough academically able students are attracted to teaching;
- . Teacher preparation programs need substantial improvement;
- . The professional working conditions of teachers are, on the whole, unacceptable;
- . The average salary of teachers is well below that of other people with equivalent training and experience; and
- . Many teachers are required to supplement their incomes with part-time and summer employment. (p. 10)

Earlier recommendations were reconfirmed by the Southern Regional Education Board (SREB) including tighter standards for admission, assessment of teacher selection techniques, a common test to aid interstate migration of teachers and a review of teacher education programs.

In an effort to determine the state of admissions into teacher education, Goertz, Ekstrom, and Coley (1984) studied the impact of state policies on teacher education. They found state-initiated policies which they call "filters" or "screens" that, in effect, control the supply of teachers. The processes used to select students entering teacher education programs included the use of college entrance examinations, basic skills tests and other criteria, such as grade point average. Prospective teachers have also been asked to pass tests in general knowledge, pedagogy and in subject matter specialization.

Goertz et al. (1984) report that the approaches used by most states are inadequate to enhance teacher quality for the following reasons:

1. Current policies filter out people rather than develop talent.
2. Restricting access reduces the socio-economic and racial-ethnic diversity of the nation's teaching force.

3. Policies focus on assessing limited number of skills which vary in degree of relevance.
4. Current policies mistakenly place responsibility on education departments instead of on the academic departments which have not taught the tested skills and knowledge.

Smith (in press) has done a comprehensive survey entitled, Minority Performance on Teacher Competency Tests: A State by State Analysis. His data collection (see Appendix C) indicates the negative impact that testing is having on black, Hispanic, and Native American teachers. He reports the following:

1. Since 1978 the number of new teachers produced by 45 predominantly black American Association of Colleges of Teacher Education (AACTE) member institutions has declined 47%.
2. In Alabama, black candidates passed 43% of the tests, compared to 86% for Anglos (1981-1983).
3. In Arizona, the pass rates were 41% for blacks, 36% for Hispanics, 25% for Asians, 19% for Native Americans, and 70% for Anglos (1983).
4. In Florida, reports for 1982-83 on the Florida Teacher Certification Examination show a 90% to 92% pass rate for white teacher candidates, 35% to 37% for black candidates, 51% to 57% for Hispanic candidates.
5. In Georgia, in 1983, with 22,261 students taking the Criterion Referenced Teacher Certification Test (CRTCT), 34% of the black candidates passed, compared with 87% of the white candidates. The 1,184 black candidates who passed the CRTCT made up only 6.7% of the teachers qualifying for certification.
6. In Louisiana, 15% of black teacher candidates compared with 78% of the white candidates have passed the National Teacher Examination (NTE). Just over 40 black students per year pass the NTE, which has reduced the number of black teachers drastically in Louisiana. Thirty-seven percent of school children and 47% of the teaching field is black. With testing, only 5% of new teachers are black based on data between 1978-82.
7. For Mississippi, competency assessment cannot be determined with exactness from available data. The pass rates for candidates from historically black institutions range from 54% to 70% compared with 97% to 100% for predominantly white state institutions.
8. Oklahoma administers a criterion-referenced and competency based test. The over-all pass rate shows black candidates, 45%; Hispanics, 71%; Native Americans, 70%; Orientals, 82%; and Anglos, 79%

- (1982, 1983, 1984 results).
9. In Texas, in the first official testing of the Pre-Professional Skills Test (PPST), passing rates for blacks was 10%; Hispanics, 19%; and whites, 62%.
  10. In Virginia, the pass rate was 56% for black candidates and 97% for white candidates (1984) on the NTE.
  11. There is a low correlation between self-reported grade point averages (GPA's) and the pass/fail status on the California Basic Educational Skills Test (CBEST) in California. This raises a serious question about measuring competence with the CBEST, not only for minority candidates but for all candidates.
  12. American Testing Service, a test designer, is not prone to publish test results by ethnicity which can make data collection difficult for researchers seeking data on minority groups.

Anrig (1985) claims the original goal of fairness and educational equality mandated by Brown vs. Board of Education of Topeka has been sidetracked. Today more than two hundred state and local task forces, panels and committees have been convened in response to the demands for excellence and decrying of mediocrity by the prestigious commissions. Imig (1984) states, "More than 300 studies, reports and analyses were completed, state governments mandated a host of reforms and innovations, and some states appropriated massive sums for the revitalization of public education" (p. 2).

Consensus exists among educational leadership at the highest national level. It is argued that there is a need to enforce a common set of entry level standards and to oppose alternative routes to certification which, in effect, deny the importance of professional knowledge and skills.

Tests and test scores have become strong political weapons. Evidence that can be cited with numerical indicators provides great support to politicians wishing to show better schools through testing. Educators have been reluctant to oppose the use of tests because they depend on the public for support. There is also a strong indication that teachers are lacking in understanding tests and application of testing. According to Salganik (1985), there is a concern on the part of educators that testing for solutions to problems in education will be viewed as being mechanistic with input-output processes which are unnatural and technical in nature with proposed solutions becoming self-defeating.

The reform movement drew the attention of the Council of

Chief State School Officers (CCSSO) to deal with the problems of attracting, preparing, licensing, and retaining teachers. Other agencies such as the American Association of Colleges of Teacher Education (AACTE) provided models for the preparation of teachers. In 1982 the National Education Association called for a reform in teacher education in its report, Excellence in Our Schools, Teacher Education: An Action Plan (1982).

A projected teacher shortage was reported by the National Center for Education Statistics in 1980 (see Appendixes D and E). Only 3% of the males who were college bound reported their intended major in education as compared with 6% in 1972. The number of females also dropped from 19% to 10% (Sikula and Roth, 1984). Stewig (1985) reported that 8% of all U.S. teachers leave the profession permanently each year and that 70% express dissatisfaction with their jobs.

According to Goertz et al. (1984), by 1988 twenty-nine states will require aspiring teachers to pass a state-prescribed, standardized test before entering a teacher education program. Nine states will require successful performance in internship programs ranging from one to three years. This will further strain teacher supply and demand.

Additionally, there is a severe shortage of prospective minority teacher candidates, other than Asians. The number of minority students enrolled in college is far below the percentage of college age minorities in the general population (see Appendixes F and G). For example,

. . . Hispanics represented 7.5 percent of all 18 to 24 year olds counted by the census bureau in 1980, but only 2.9 percent of the college population that year. Similarly, the black college-age population of 13 percent translates to 8.4 percent of college enrollments, and Native Americans were 0.5 percent of the college population that year compared to 0.7 percent of the population in the age group.

Conversely, the white and Asian college enrollment percentages exceeded their percents of population for 18 to 24 year olds. Whites made up 80.9 percent of the age group and represented 82.9 percent of college students, and Asian teens were 1.5 percent of the age group and 2.1 percent of the college population.

If the under-representation trends continue, the future looks bleak for Hispanics, blacks, and Native Americans. ("Minority Education Trends," 1985, p. 45)

The following critical trends substantiate the rapid

growth of minority populations and the reduction in minority graduates, ("Minority Education Trends," 1985):

1. The U.S. Department of Commerce found that 29.9 percent of Hispanic high school grads in 1980 continued to college. Five years earlier, 35.4 percent of the Hispanic graduates had entered a college or university.
2. Among high school freshman, less than one-half the Hispanics and blacks made it to graduation day, according to a New York State Department of Education study on the class of '83.
3. The MHE [Minorities in Higher Education] study briefly discusses minorities in academe, first stating that "there are very few minority academics"-a scant 3 percent in 1979.
4. Of the 32,839 doctor's degrees conferred in 1981, 454 went to Hispanics or 1.4 percent.
5. It found that of the 399 schools that awarded doctorates in 1981, 269 awarded no doctor's degrees to Hispanics and 211 awarded no doctor's degrees to blacks. Heavy regional concentrations of Hispanics and blacks may account for the fact that 59 percent of the Hispanics receiving doctorates, received them from 8 percent of the doctorate-granting institutions; or that 60 percent of the blacks conferred doctorates were conferred by 10 percent of the schools conferring doctorates; but, lack of recruitment, limited financial assistance and restrictive admissions criteria are "equally compelling reasons" for the skewed numbers, states the report.
6. The country's "prosperity and well-being" may well depend of the educational accomplishments of the projected 35 percent minority population in the year 2020. (pp. 45-46)

#### Teacher Competence

With the reform movement clearly mandating tests for prospective teachers, states are no longer willing to accept teachers with degrees as guaranteed competent professionals (Benderson, 1982). The problems of teacher competence go beyond the proliferation of tests. With the woman's movement, many competent females continue to select other higher paying professions over teaching.

Short (1985) states that the term "competence" is one of the most misused and overused concepts in the reform movement. The ability of a teacher to pass a basic skills test or to perform a highly developed teaching skill is often referred to as teacher competence. Competence and

incompetence are often referred to as being distinctly identifiable like black and white. Politicians, the media, the public and even many educators talk about competency as being clearly defined and measurable.

Smith (in press) states that a number of observations can be made about the teacher competency testing movement:

First, the initiative for the movement came from state legislatures and boards of education rather than from teacher organizations or college and university schools and departments of education, thereby weakening the position of educators to govern their profession. Second, in several states the testing of teachers is paralleled by a movement to assess the basic skills of public school students with the latter, more often than not, having been mandated first. Third, competency testing of teachers has taken root despite inadequate research to show a direct relationship between performance on pencil-paper tests and on-the-job competence. Fourth, recognizing that predictive validity of the tests has been generally unsubstantiated, most state departments of education and test developers have been cautious to establish content validity through textbook review and through test item review by teachers and other groups of professional educators. Fifth, the tests are more often norm-referenced rather than criterion-referenced, a fact that makes designing successful college developmental and teacher preparation programs extremely difficult. Finally, state determined cut scores have been consistently set at a level high enough to eliminate disproportionate numbers of minority students from the teaching profession. (p. 9)

Short (1985) believes that competence should be defined in terms of precise and measurable criteria. One aspect of competency should be the measuring of knowledge and skills necessary to perform on the job. Another aspect should be the determining of the level of competence at which the individual must perform. Identifying and measuring all aspects of competence should then be established holistically in order to measure the entire person's competence (Short, 1985).

Bridges and Groves (1984) studied dismissal for incompetence of tenured teachers over a forty-three year period. Their research shows a clear difference between the content of tests being used to filter prospective teachers from entering teacher education programs and evaluation for teacher retention. Most teachers are dismissed for the following reasons:

1. Technical Failure. The teacher's expertise falls short of what the task requires.
2. Bureaucratic failure. The teacher fails to comply with school/district rules and regulations or directives of superiors.
3. Ethical failure. The teacher fails to conform to standards of conduct presumably applicable to members of the teaching profession.
4. Productive failure. The teacher fails to obtain certain desirable results in the classroom.
5. Personal failure. The teacher lacks certain cognitive, affective, or physical attributes deemed instrumental in teaching. Indicators of personal deficiencies include poor judgment, emotional instability, lack of self-control, and insufficient strength to withstand the rigors of teaching.  
(Bridges & Groves, 1984, pp. 6-7)

Bridges and Groves (1984) also state:

Criteria which have appeared in previous dismissal cases heard at the appellate court level are as follows:

1. knowledge of subject matter
2. ability to impart knowledge effectively
3. ability to obtain the respect of parents and students
4. proper use of corporal punishment
5. willingness to accept teaching advice from superiors
6. adequate academic progress of students
7. ability to maintain discipline
8. physical ability to perform the duties of a teacher
9. emotional stability

The following types of information may be used in evaluating teachers:

1. Supervisor ratings
2. Student ratings
3. Student performance on tests
4. Peer evaluations
5. Self-evaluations
6. Student and parent complaints (pp. 16, 20)

The administration of teacher competency tests is likely to bring about contract terminations without regard for proven teacher effectiveness. Kelleher (1985) describes the incompetent teacher as one "who has demonstrated his or her inability to meet minimum standards of performance over a number of years" (p. 362). Teachers need an extended period to demonstrate competence with measurable behaviors as well as other validated criteria.

The question is, what is teaching effectiveness and can it be measured accurately. The complex teaching process can be described as a practical art which requires cognitive awareness, practice and dedication. Gage (1984) describes teaching as follows:

As an instrumental art, teaching departs from recipes, formulas, and algorithms. It requires improvisation, spontaneity, the handling of a vast array of considerations of form, style, pace, rhythm, and appropriateness in ways so complex that even computers must lose the way, just as they cannot achieve what a mother does with a 5-year-old.

The sense in which I use the term art includes any process or procedure whose tremendous complexity - resulting from the large number of relevant variables and the interactions among those variables - makes the process irreducible to systematic formulas. (p. 88)

Wisniewski (1984) states:

If we avoid dealing with personality variables, we concurrently miss the rich social dynamics that are the heart of teaching. Teaching is essentially a balancing act involving teachers and students. Whatever the subject or grade level, the balance between expectations and behaviors include give and take, command and response, goad and reaction, fear and praise, respect and hate, participation and withdrawal, passion and boredom, and every other characteristic of human interaction. (p. 36)

Wisniewski indicates that it is much "safer" to try to reduce teaching into standardized measurements thus avoiding the subjective judgments required by humans. Thus, the prominence of testing in our society has made the unquestioned use of tests and test scores "acceptable."

#### What Tests Measure

Predictive validity of a test rests with its ability to show a significant relationship between a test score and actual performance. Stedman (1984) criticizes test developers and examiners for disregarding the importance of the predictive validity of tests:

The true value of tests lies in their predictive power or their ability to estimate success in whatever it is they are to measure. A study of scores from American College Testing (ACT) reported by Monday and Davis (1974) indicated that there was no relationship between adult accomplishment and academic talent. The study

further implied that accomplishment in leadership, music, literature, art, drama and science is uncorrelated with high school or college GPAs and certain test scores.

These data are not at all unusual or unexpected. A number of previous studies have been conducted to study the effectiveness of grading procedures at the college level. Pallett (1965) found no correlation between college grades and ratings on eight dimensions which characterize success in business. It was similarly found that academic success was independent of up to 26 performance characteristics of physicians (Prince, Taylor, Richards, & Jacobsen, 1963). Also, a review of 33 studies indicated that grades had little correlation with performance success when GPAs were compared to a supervisor's rating in teaching (Hoyt, 1965).

Predictive validity relates to the accuracy of determining who is likely to succeed or fail in professional practices based upon data collected from some form of evaluations (i.e., a test). It may also be applied to establishing evidence that a given program or set of experiences will produce necessary skills to improve practice. Unfortunately, few professions have been successful in establishing predictive validity for their basic programs, entry exams, licensing exams, or continuing education activities. There is some concern that current psychometric theory is inept at providing guidance for such determinations.

Too often tests are selected for purposes which are incompatible with their original design, while others are interpreted in improper ways. In Tennessee, for example, the California Achievement Test was adopted as a measurement of facility in basic skills for entry into teacher education. No effort is underway to evaluate whether this test is adequate for its intended purpose. No data exists to support either the choice of test or the level of competency required regarding potential for success in teacher education programs. (pp. 2-3)

There are many problems involved in the correlation and predictive power of tests. Educational pencil-and-paper tests present the examinee with primarily cognitive tasks and have no predictive value. "Educational tests are all, in one way or another, tests of cognitive development" (Ebel, 1977, p. 29). Standardized tests favor native-born white American children (Williams, 1970). The two main problems of standardized testing involve scientific and ethical decisions. Standardized tests are inadequate measures of the capabilities of minorities (Messick & Anderson, 1970). In 1975, the idea that a culture free test of general learning ability would be practical had simply not been developed (Young, 1975). According to

Woolever (1985), standardized rating forms used for teacher evaluation are assumed to measure elements of quality teaching and to be reliable. "This is clearly not the case" (p. 23). The accurate measurement of teaching performance is beyond our current ability to measure (Soar, Medley, & Coker, 1983; Sykes, 1983; Zahorik, 1984).

Rosenbach (1979) believes that most claims of test bias are unfounded, especially in intelligence and aptitude tests. He also says that tests are accurate predictors of school achievement but not very predictive of success ou. side of school. Hoover (1984) argues that the Florida Teacher Certification Examination is culturally, linguistically, and philosophically biased against minorities.

School administrators in areas experiencing teacher shortages are pleading for more study on competency tests and greater input from public school teachers and administrators (Woods, 1985).

Kenneth White, superintendent of Mission Independent School District in Texas, discussed competency testing for teachers with Governor Mark White of Texas and pleaded for more study and greater input from school people. He stated, "We also asked for safeguards of present certification and extended time to reform" (Woods, 1985).

Gideonse (1985) claims that tests do not improve the caliber of those entering the teaching profession. Tests can only measure what people know, not how they can perform. Examination use denies access to minimally qualified prospective teachers and is unjustifiable. Gideonse predicts that current testing will bring few lasting benefits and attributes its existence to the vulnerability of the teaching profession.

There is some evidence that learning to take tests can improve test scores. Black high school students showed more improvement on the Scholastic Aptitude Test than did whites ("SAT Scores," 1985). Unfamiliarity with examinations and inaccurate interpretation of test questions often have a negative impact on minorities. However, little evidence exists that learning to take tests helps elevate scores significantly, according to Gregory Anrig, president of Educational Testing Service (Bowen & Booth, 1985).

In his study of academic performance of black and white students on a commuter campus, Mannan (1985) states that standardized tests are poor predictors of black performance. Affirmative action gains made by blacks (Evans, 1985) are found to be eroding because of the testing movement. The

disproportionate numbers of minority candidates in teaching continues to reflect the underrepresentative characteristic of early historical periods.

The predictive value of tests is questionable because the reasons needed to explain the poor performance of minorities have not yet been given (Raspberry, 1985). The literature is filled with concerns about fair testing of ethnic minority groups. Studies cover the areas of test validity, test bias, racial differences, equal job opportunities, tests' predictive validity, school admission, counseling, and cultural factors. The general use of tests is harmful to populations that do not fit the mainstream of society through misclassification, test unfairness, and questionable accountability (ERIC Highlights, 1980).

Smith (in press) refers to testing and its impact on minorities as follows:

Today we chuckle at [tests'] absurdity to mask our shame. In future years our shame may be no less as we look back on competency testing.

Clearly, any professional practice that excludes disproportionate numbers of minorities represents neither excellence nor equity. If this nation is considered at risk now, a decade of willful elimination of minority teachers will result in a nation lost. (p. 43)

Test development has come under close scrutiny by researchers. One of the major problems in the use of educational and psychological tests is the need to define scales in terms of some specific reference group. Scaling is used to protect against bias when the composition of groups vary. Even though there have been no legal challenges of these procedures to date, the users have the responsibility to practice only fair, accurate, and equitable tests (Donlon, 1985).

Much of the literature focuses on standard setting of norm referenced tests using the processes developed by Angoff and Ebel (cited in Holmes, 1985). Each time a new form of a criterion-referenced test is developed, it must go through a standard-setting study--a process that is prohibitively expensive and time consuming. The techniques used in determining a performance standard for a test usually call for a percentage correct score. Criterion-referenced tests are based on absolute, rather than relative, decisions about test content and individual performance and place great importance on specific items. It is very disappointing to see the different results reported from the same technique applied to the same test

(Andrew & Hecht, 1976; Halpin, Sigmon, & Halpin, 1983; Mills & Barr, 1983; Saunders, Ryan, & Huynh, 1981; as cited in Holmes, 1985). The probability of gauging performance with consistent accuracy through existing methods is very unlikely (Garcia, 1971). According to Holmes (1985), improper test use often presents serious harm to the public and to individuals who have invested so much of their time, resources, and efforts in becoming educators.

The various California Basic Educational Skills Tests (CBEST) are "scaled" to achieve comparability. The accuracy of equating results will depend on the degree to which two tests differ. Test equating is appropriate only when forms of the test equated measure the same trait or attribute. The new Standards for Educational and Psychological Testing (1985) call for stricter guidelines insuring that those tests being equated are parallel. It is more important to construct tests carefully than to try to compensate for the results of poorly constructed tests later on. No single method used has been identified by research as being most accurate (Holmes, 1985). Current methods used in test development and application raise questions about the validity of such tests and may help explain the varying scores reported by individuals retaking the same test.

Cross, Impara, Frany, and Jaeger (1984) compared three methods used to establish minimum standards on the NTE as a basis for content validation. Of the Angoff, Nedelsky, and Jaeger procedures used, only the Jaeger method requires normative data and was favored over the other two. No studies could be located that compared methods across tests of differing content.

The courts' role in affirming the content validity of the National Teacher Examinations (NTE) has been pointed out by A. J. Wilson (1984b) who states:

By the 1970's, renewed attention was paid to questions of test content and test validity with the emergence of a number of law suits charging that the tests were being used in some states and communities to discriminate against minority teachers and teacher candidates. In 1978, U.S. Supreme Court, relying heavily on a study conducted for the state of South Carolina by the Educational Testing Service, ruled in favor of the exams' use and thus indirectly in favor of their content validity. The court affirmed a lower court's decision which stated that "the NTE have been demonstrated to provide a useful measure of the extent to which prospective teachers have mastered the content of their teacher training program." (p. 2)

Yalow and Collins (1985) state, "The processes that can be used to help ensure that these assessment procedures, often paper-and-pencil tests, possess content validity and, subsequently, to demonstrate this fact, are not yet standardized" (p. 3). Various processes have been used by test developers to successfully stand off legal challenges regarding content validity of tests. Test items are reviewed by a select population of experts who are asked for an opinion on a series of given skills and their applicability to acceptable performance. A belief that teachers should be able to do something is not sufficiently powerful for inclusion as a skill. The skills/knowledge assessed must be necessary for job performance in order to hold up under scrutiny of content validity.

Another method used to demonstrate content validity is the "51 percent rule" used by ETS in South Carolina, North Carolina, and Mississippi. If over 50% of respondents indicate that an item is valid, it is considered to be content valid (Yalow & Collins, 1985). However, since employers cannot justify employment tests on the grounds of content validity if they cannot prove that the content of the examinations includes critical and substantial parts of the job, this procedure is questionable. Any one method used to substantiate content validity of a test is not defensible. There should be an accumulation of sources in order to make a claim of content validity more or less plausible. If a test is given as a screening device, it must reflect knowledge learned, such as that learned in a teacher education program (Yalow & Collins, 1985).

#### Major Tests in Use

Before discussing specific tests, the cooperation received from the Educational Testing Service in gathering data for this study must be acknowledged and commended. Since ETS is the largest test developer in the nation, a major part of test analysis refers to its work. Test developers like ETS do not take responsibility for test validation, cut-off scores, or test use. ETS does on occasion, however, carry out validation studies on behalf of those states requesting them. Test developers design examinations based on expert advice and, in most cases, from existing test files. The demands for new or different tests has strained the capacity of most test developers to produce the types of tests being mandated by state legislatures, political entities, and the public. Consequently, examinations with very similar test specifications and testing the same domains have appeared commercially. However, they are being used in different ways, with different cut-off scores, for different purposes.

## The National Teacher Examination

The National Teacher Examination (NTE) now administered by the Educational Testing Service (ETS) has a history extending over forty years. It is the most widely used examination program (Benderson, 1982). Developed originally to assist superintendents in assessing teacher knowledge for hiring purposes, the test had two parts - specific "teaching area examinations" and "common examinations" for all teachers (A. J. Wilson, 1985). In the early 1960's, the test took on greater use for determining certification of teachers and salary increases. In 1979 ETS named a "NTE Policy Council" to establish program policy on test development administration, and use. In 1982, the "NTE Core" replaced the "common examinations" and the "area examinations" became "special area tests" (A. J. Wilson, 1984a).

The NTE Core was designed to test content related to both teacher education curricula and existing requirements into entry level teaching positions. It was designed to measure achievement in the liberal arts and professional components of training common to most teacher education programs (Rosner & Grandy, 1982). There is reason to wonder how accurately measurement may be reflected since teacher education requirements vary considerably from state to state and from institution to institution.

The NTE Core Battery Tests are standardized examinations that provide objective measures of knowledge in three major areas as follows:

- A. Test of Communication Skills
  - 1. Listening
  - 2. Reading
  - 3. Writing: Multiple-Choice
  - 4. Writing: Essay
- B. Test of General Knowledge
  - 1. Literature
  - 2. Mathematics
  - 3. Science
  - 4. Social Studies
- C. Test of Professional Knowledge
  - 1. Three 35-question operational sections
  - 2. One 35-question pretest section (ETS, 1982)

Each of the three tests is two hours long and divided into four thirty minute sections. All tests consist of multiple choice questions or problems except one portion on communication skills which requires an essay that is scored holistically.

According to the test developers, the professional knowledge test has as its underlying base the conceptualization of the beginning teacher as an intelligent, informed decision-maker. The test items are based on knowledge relevant to the content and process of teaching which the beginning teacher is expected to have. The material is determined to be at the appropriate level of difficulty of a beginning teacher with up to two years of teaching experience (ETS, 1982). It should be noted that the NTE is being used for determining entry into teacher education programs, exiting programs, certification, and for hiring purposes.

The NTE is a standardized examination in which validity relates most directly to the content of teacher education programs. The broad scope of the knowledge tested indicates that the best preparation for the test is the knowledge and experience one gains from a teacher preparation program. The scores are not exact measures, and each score has associated with it a degree of error. NTE policy does not set passing scores for the examinations. ETS offers guidelines for setting minimum standards and test score use. It encourages recipients to use test scores as one of several bases for making decisions about certification and selection of teachers (ETS, 1985). However, Toch (1984) reports an increase in the use of the NTE as an admissions requirement. This has reduced the number of blacks entering the teaching profession.

Olstad, Bean, Foster, and Marrett (1985) did a validity study on the general knowledge component of the NTE Core. They found a rather consistent correlation between the NTE Test of General Knowledge and the California Achievement Test (CAT). They also note a lack of correlation between the GPA and either the NTE or the CAT which suggests that they measure different factors. A 1983 study by Olstad, Beal, Schlick-Noe, and Schaefer at the University of Washington suggested a predictive value of GPA for teaching performance (cited in Olstad et al., 1985), whereas, the CAT did not show a predictive value. According to Olstad et al., the correlation shown between the NTE and CAT would indicate that the NTE has no predictive value in assessing success in a student teaching program. The NTE has not been empirically correlated with student teaching outcomes. Peterson (1984) cites a correlation reported in Georgia and Oklahoma of .30 and .40 respectively between student teaching and on-the-job performance.

#### The Pre-Professional Skills Test

The development of the Pre-Professional Skills Test (PPST) took place over a relatively short period of time due

to the high demand for examinations placed on ETS. The test specifications from the NTE and those of the PPST are nearly identical. M. Goodison of ETS, Princeton (personal communication, May 16, 1985) states:

The PPST specifications are a subset of the Core Battery specifications and were not developed independently. Though the items in the two tests measure the same skills, they are a different set of items, and other than in the first form of the PPST, do not overlap with the Core Battery Tests.

The PPST is not "normed." There is no data base on a national representative sample. The meaningfulness of the score is not dependent on scores from a national sample. This type of test measures what a student can do without comparison with other populations. The proposition is made that these tests should be thought of as criterion-referenced tests (Goodison, 1985).

The PPST is most often used for admission into teacher training programs. Some states use the test scores for counseling and remediation purposes. Three states--Texas, Tennessee, and Kansas--require students to pass the test with a given cut-off score while Delaware, West Virginia, Nevada, and New Hampshire presently or will soon require the PPST for initial certification. Some school districts are also using the test for teacher selection (Goodison, 1985).

It is impossible to "measure" individual knowledge or effective teacher performance. Placing such demands on testing systems rather than on performance is impractical and unfair (Kidd, 1985).

#### The California Basic Educational Skills Test

The CBEST is almost a "clone" of the PPST. It was designed from the PPST for the purpose of providing a general measure of basic proficiency in reading, writing and mathematics. The scores are intended to provide only an overall indication of the examinees' strengths or weaknesses in each of three skill areas. The main difference between the PPST and the CBEST is that the PPST has a writing component composed of an objective (multiple choice) section and a written essay, whereas, the CBEST has two written essays. There is an indication that the reliability of the CBEST may have been weakened by the substitution of the written essay for the objective writing portion (Peterson, 1984). The written portions are scored holistically.

Wright (1985) reflects a thoughtful concern for the CBEST's potential effect on teacher supply, especially as it

relates to certain ethnic minority groups. In this assessment, the CBEST is generally judged to be relevant to tasks that entry level professional staff members must carry out.

Generally, the questions [of the CBEST] are judged to be easy or medium in difficulty for those who are required to take the test. The scoring standards applied for the writing section of the test are generally consistent with the standards that school practitioners would apply.

As noted, however, there are substantial portions of the review panel members who would disagree quite strongly with the substance of the preceding paragraph. If changes in the test were to be considered, the following would be advocated by these people;

1. Increase the weight given to literal comprehension reading skills, and decrease the weight given to critical comprehension.
2. Increase the weight given to mathematics questions based on arithmetic, and decrease the weight given to algebra questions. Also, to the extent possible, decrease the amount of reading required for the mathematics section.

The CBEST review panel also showed a wide range of opinion concerning the passing scores on the test, but as a group, they do not think the required passing scores on the reading and mathematics sections should be lowered. (Wright, 1985, p. 41)

Table 1 is indicative of the concern of passing rates for the CBEST.

Table 1

Percent of Examinees Passing the CBEST in the First Year of Administration

(Source: Goertz, Ekstrom, & Coley, 1984, p. 130, Table 1.)

Ethnic Group	<u>First-time Test Takers</u>		<u>Repeaters</u>	
	Number of Examinees	Percent Passing	Number of Examinees	Percent Passing
White	24,540	76%	2,493	46%
Black	2,040	26	560	19
Oriental or Asian American	1,259	50	360	26
Mexican American	2,133	39	781	28
Other Hispanic	851	38	280	24
Other	1,216	63	184	35

The Scholastic Aptitude Test

Declining Scholastic Aptitude Test (SAT) scores are often cited with regard to teacher candidates. SAT scores are valid predictors only for the first year of college. The first year of academic performance does not predict performance for the fourth year because of the adjustments made during this period. Most college faculties prefer that students also possess such other characteristics as honesty and open-mindedness and be altruistic and highly motivated.

The credibility of the testing movement in the name of educational accountability is adamantly criticized by journalist David Owen in his book, None of the Above: Behind the Myth of Scholastic Aptitude (1985). He points out that Carl Campbell Brigham, the creator of the SAT, was known for his studies on the inferiority of blacks.

Owen (1985) charges that the SAT is based on a White, upper-middle class, suburban point of view, thus penalizing Blacks and other urban minorities. He indicates that cram-coaching can improve scores significantly. He charges ETS with using distractors in questions make the test deceptive. ETS president Gregory Anrig disputes the test score gains and other claims reported by Owen.

Owen claims that ETS controls who gets ahead in this country through the sale of its five hundred different tests. He suggests that ETS perpetuates the very inequities it claims to eliminate by its testing motives. His book attacks the integrity of ETS, calling for the abolition of both ETS and the SAT (Biemiller, 1985).

The low SAT scores, that have been cited continuously in the report A Nation at Risk: An Imperative for Educational Reform (1983) as an indication of mediocre public education, are challenged as having little to do with the quality of states' educational systems. Percentages of students taking the test ranging from 4% in Arkansas to 59% in New York of students are strong indicators that populations taking the SAT vary too much to draw accurate conclusions about the quality of individual and state educational programs.

The implication that SAT scores have a lesser predictability for Blacks than does GPA is clearly noted in Owen's study. The predictability of SAT scores for college performance is enhanced when used with GPA. True coaching can raise test scores without affecting criterion performance, however, and would invalidate test scores. The question of how much scores can be elevated is still unresolved between ETS's president Gregory Anrig and David Owen (Bowen & Booth, 1985).

The argument that SAT scores used to alert the nation about a "nation at risk" has been described by a number of writers as a misinterpretation of what the falling scores really mean. Composition of the population groups taking the test changed during the historical period of affirmative action. Test-taking populations changed, giving the impression that the schools were producing poorer quality students. The College Board Advisory Panel report indicates erroneous interpretations of SAT scores due to variance in population groups tested. The sex and socioeconomic backgrounds of test takers powerfully influence their performance. As test scores began to rise through the process of pre-selection of test takers, politicians will hasten to take credit for the apparent improvement of the academic delivery system (Howe, 1985).

The SAT has verbal and math components that are indicators used for admission purposes. However, as reported by Schaffner (1985), high school records play a more important role in admissions nationally than do standardized test scores.

Lewis (1985) states that the continued use of the SAT and ACT will continue to perpetuate unfair comparisons. (See Appendix H for comparative data from North Carolina.)

Ethnicity and low socioeconomic status are high predictors of failure rate. However, since test scores seem easy for the public to "understand" and since the media readily provides them, this obstacle will continue. National testing had received very little support until the tremendous diversity of standardized tests with different applications evolved. It is possible that the current national momentum may make a national teacher test a reality.

## Use of Tests

### Teacher Test Use

Tests like the NTE, CBEST, and the PPST measure knowledge and certain abilities such as reading, writing and math. They do not measure teacher performance, classroom control, motivation techniques, application of knowledge in a teaching situation, personality, or stress. Tests such as the NTE were never intended to predict teaching performance, and there are very low correlations between measures of teacher effectiveness and test scores (Peterson, 1984).

Stoker and Tarrab's (1985) study on the correlation between the American College Test (ACT) and the PPST revealed the composite ACT correlated with all three PPST scores and highest with the total PPST score. They suggest that if future studies show these same results, it may make the PPST unnecessary due to the long established reliability of the ACT. They also report the distressing possibility that college English and math courses may not improve students' abilities to score well on the PPST, which is verified by ETS (Goodison, 1985). Stoker and Tarrab's recommendation is to use the ACT for admission into teacher education programs. The ACT has also been very useful in predicting the Medical California Achievement Test (MCAT) subtest scores (Dawson-Saunders, Paiva, & Doolen, 1985).

The diversity of uses for the PPST in various states, with each state determining its use and its own cut-off scores, raises ethical questions about probable misuse of the test which is inconsistent with ETS guidelines and professional practices as cited in the Standards for Educational and Psychological Testing (1985). Current practices in content validation of tests vary considerably from legal court decisions or settlements. Questions of test validation must go beyond consideration of content validity.

In gathering information, the principal investigator interviewed Marlene Goodison, Program Administrator for the PPST at Educational Testing Service and had various

interviews with other ETS specialists at Berkeley, California and Princeton, New Jersey, including a brief conversation with Dr. Gregory Anrig, president of ETS. The following conclusions are made based on these communications:

1. ETS put together the first PPST test on short notice from existing question files and the NTE Core Battery.
2. ETS warns that the use of the PPST as sole criteria for any purpose such as entry into teacher education programs should be strictly avoided.
3. States are responsible for the appropriate use of the PPST. (The courts have dealt with the issue of test developers having responsibility for test use.)
4. There are substantial differences between the pass/fail rates of ethnic groups or races on the PPST. Hispanics and Blacks score significantly lower than Whites in all three tests, especially in the reading and mathematics portions.
5. Remediation designed to significantly increase PPST test scores to ensure pass rates does not exist.
6. Those who have missed the qualifying score by a few points have a reasonable chance of improving their scores enough to pass a second attempt. However, large score gains are rare.
7. Generalizations made on scores should be made only for PPST first-time test takers (Goodison, 1985).
8. Male testers score higher than females in mathematics while females score higher than males in writing.
9. In all cases, use of the PPST should be supported by local content validation studies, making the users responsible for use.

In Texas, the PPST may be taken an unlimited number of times to achieve a passing score. Fisk (1984) offers optimism on PPST retakes and suggests that students may improve their test scores. However, ETS warns against predicting retake gains for scores other than the first-time test takers. Reported percentage increases in retake examinations in Texas have not been made on original populations. The percentage increases reported in Texas may have come about through pre-selection of students and good advisement.

Use of standardized examinations may violate the principal of equity, an important aspect of the democratic way of life in America. Standardized tests have been designed for predominantly white populations with distinct linguistic styles different from those of minority populations. However, some population groups score higher on specific subtests. For example, Jewish students tend to

score higher on Scholastic Aptitude Test (SAT) verbal ability and Chinese-American students on SAT spatial conceptualization (Willie, 1985).

Some institutions with low minority enrollments, such as Texas A&M with 500 Hispanic students on a campus with 39,000 will not feel much of an impact from the PPST. On the other hand, institutions with open admissions having a large minority population will be required to make major program and fiscal modifications if student population changes occur. The negative impact reported by Smith (in press) on minorities is likely to happen in Texas.

Goodison (1985) states that on the PPST:

Substantial differences in pass rate by race or ethnic group occur in all three skill areas. Blacks, Mexican Americans, and other Hispanics generally show somewhat less than or approximately half the pass rate of Whites, in Reading and Mathematics. In Writing, the minority pass rate ranges from about half the White pass rate to about two-thirds the White pass rate.

Data on PPST repeat test takers accumulated from February 1983 through December 1984. . . . Approximately 40 percent of the Reading test repeaters showed no score gain. The most frequently observed score gain in Reading is four score points. About a third had score gains of one to four points. Less than four percent increased their scores by 10 points or more. On the Mathematics test, about one-third of the repeaters showed no gain. The most frequent gain was three score points. Nearly 40 percent of the repeaters score gains were from one to four points. Fewer than 4 percent improved their scores by ten points or more. The Writing test data . . . show that twenty-nine percent of the repeaters had no improvement on the retest. The most frequent score gain was two points; nearly half gained one to four points. Fewer than 2 percent improved their scores by ten points or more.

The data on repeaters suggest that examinees who have missed the qualifying score by only a few points have a reasonable chance of improving their scores enough to pass on a second attempt. However, large score gains are rare (pp. 12-13).

There is a great diversity in test use and cut-off scores for teaching credentials and certification. California uses the CBEST for entry into teaching and service credentials. Florida requires teachers seeking certification to pass the Florida Teacher Certification Examination (FTCE). Georgia has required the Teacher Certification Test (TAT) since 1978 in the field of

certification. In Louisiana, applicants for teacher certification are required to pass the Core Battery (the Commons Exam prior to 1982) and the specific test in their area of specialization (Carlson, 1985).

### Misuse of Tests

Educators who must be evaluated by them are concerned about tests that are used and their interpretation. A recent survey showed that 97% of those teachers who took a trial test indicated the test did not measure their ability to teach ("Protesting the Tests," 1985). Claims that competency tests do not measure teachers' abilities or talents (Stolarek, 1985) are being proclaimed at a time when indicators of a teacher shortage appear on the horizon. As some states develop examinations, other states adopt them for use, i.e., the Georgia test received approval in Arizona and Mississippi (Fanner & Ebers, 1985).

The development and appropriate validation of defensible tests can be very time consuming and costly. According to The Standards for Development and Use of Educational and Psychological Tests (1985), appropriate test validation requires an extended time period. It is important to realize that validity is itself inferred, not measured. There are four independent kinds of inferential interpretation which traditionally describe validity related to most professional test use. There are criterion related validities (predictive and concurrent), content validity, and construct validity. The mere appearance of validity, often referred to as "face validity," is not an acceptable basis for interpretive inferences from test scores. Most validation studies on tests used for program entrance, for credentialing, or for certification do not adequately meet all of the validation criteria they should. Test developers are not required and tend not to attempt to meet these standards because the courts, (e.g. U.S. v. South Carolina, 1978), have accepted more narrow interpretations of validity.

The validation process of the NTE Core Battery and Specialty Area tests done by ETS for the New York State Department of Education and the validation study performed for the Tennessee State Department of Education by the Bureau of Educational Research and Services, College of Education at Memphis State University, essentially follow the format used by ETS in earlier validation studies. Using the same format recommended by a test developer does not insure that the appropriate validation for a specific use has been met.

## Guidelines for Proper Test Use

Test developers have continued to warn about inappropriate test use. ETS has established clear procedures for use of its tests and believes that "proper and fair use of ETS tests is essential to the social utility and professional acceptance of ETS work" (ETS, 1983, p. 21). The Guidelines for Proper Use of NTE Tests, published by ETS (National Policy Council, 1983), clearly identifies the NTE tests as follows:

- . the Core Battery, which measures areas common to most teacher education programs;
- . twenty-seven Specialty Area tests, designed to assess examinees' preparation in specific subject fields;
- . the Pre-Professional Skills Tests (PPST), which measure basic proficiency in reading, writing, and mathematics;
- . tests developed under special contract. (p. 3)

The appropriate use of NTE tests places direct responsibility on the user as indicated in the guidelines:

Sound professional practice requires that any such test be validated through appropriate studies by the test user for the specific purpose for which it is being used. Thus, the importance of validity studies carefully conducted by states, institutions of higher education, school districts or other agencies using NTE tests cannot be overstated. (p. 5)

For state agencies responsible for credentialing, ETS indicates that the proper use of tests requires the following:

When NTE tests are used for determining eligibility for certification, the certifying agency should:

- a. Ensure that multiple criteria, such as those required by relevant laws and/or regulations, are employed for certification;
- b. Publicly promulgate those criteria;
- c. Involve constituent groups, such as teacher educators, classroom teachers, and school administrators, in the process;
- d. Validate the tests to determine that they measure a representative sample of the knowledge and skills required for certification of beginning teachers, and that any qualifying scores used differentiate between those who do and those who do not possess the requisite knowledge and skills. Such validation should conform to applicable professional standards and

to federal, state, and local laws and regulations. (p. 6)

When NTE tests are used by school districts, ETS provides the guidelines for appropriate use which includes the use of multiple criteria such as interviews, references, test performance, classroom observation, written applications and academic transcripts. The guidelines also state:

The current NTE tests were developed to provide information about candidates' academic knowledge and skills, typically acquired through a teacher-training program. They do not provide a direct evaluation of teaching performance. For this reason, NTE tests should not be used by school districts, directly or indirectly, to determine the compensation, retention, termination, advancement, pay supplements, or change in provisional employment status of teachers once they are employed. Such decisions about teachers in service should be based on teaching competencies as determined directly by the supervisory and evaluation procedures of the employing school district.

The current NTE tests measure knowledge and skills needed by the beginning teacher; more is required of the teacher in service. . . . If current NTE tests are to be used as part of a program of continuing professional education, they should not be used, either directly or indirectly, for determining the compensation, retention, termination, advancement, pay supplements, or change in provisional employment status of teachers once they are employed. (p. 7)

For use by institutions of higher learning and use by college, universities and state governing boards for public higher education, the appropriate guidelines are:

When NTE tests, such as the Pre-Professional Skills Tests (PPST) or portions of the Core Battery, are used to evaluate candidates for admission to teacher-training programs at colleges or universities, the college, university, or state agency should:

- a. Ensure that multiple criteria are employed for selection;
- b. Validate the NTE tests to determine that they measure knowledge and skills important to successful completion of and graduation from the teacher-training program. Such validation should conform to applicable professional standards and to federal, state, and local laws and regulations. (p. 8)

The validation process rests with the user of the tests.

ETS clearly states the appropriate use of its tests as well as other tests as follows:

Sound professional practice requires that NTE tests be validated for the specific purposes for which they are being used. In addition, federal and other civil rights laws, such as Title VI and Title VII of the Civil Rights Act of 1964, may also require validation if the use being made of the tests is shown to disproportionately discriminate against disadvantaged members of ethnic, racial, religious, or gender subgroups.

1. Validity studies should establish that the tests are a valid measure of the knowledge and skills required for the intended use. Whether conducted by ETS or others, the studies should be designed to comply with professional and legal standards, such as the Standards for Educational and Psychological Tests published in 1974 and 1985 by the American Psychological Association, and controlling statutory law, regulations, and court decisions. Test users should also refer to the federal Uniform Guidelines on Employee Selection Procedures in evaluating compliance with federal law requirements in the context of teacher certification or teacher selection. In some cases, these standards require the use of job analyses or other similar techniques.
2. If the use of a test involves a cut score, the validity study should show that score to be consistent with a reasonable judgment of the minimum knowledge and skills necessary for the position. If used at all, rank ordering should be based on multiple criteria rather than based solely on test scores. Test scores should be used as part of any rank ordering procedure only when a score user can demonstrate, through a job analysis or otherwise, that a higher score on a content-validated selection procedure is likely to result in better job performance. . . .
- [3.] When validity studies for NTE tests are conducted by agencies other than ETS, ETS will, on behalf of the NTE Policy Council and at the request of the requiring jurisdiction, be available to review the study design and comment informally on areas in which ETS might suggest different methodologies, as appropriate. (pp. 8-9)

Clearly labeled as misuse by ETS are the following:

1. Using NTE tests as the sole criterion for decisions, whether concerning certification, selection, admission, or program evaluation.

2. Using NTE tests, directly or indirectly, to determine the compensation, retention, termination, advancement, pay supplements, or change in provisional employment status of teachers once they are employed.
3. Using NTE tests, whether for certification, selection, or admission without appropriate validation, or using qualifying scores without having conducted appropriate standard-setting studies. (p. 10)

The testing of educators is likely to continue unabated for quite some time. In an attempt to hastily attain "educational excellence" in America, teacher testing has acquired national momentum. The realization that there should be precautions taken in the nation's rush to mandate quality are identified by Anrig (1985) as follows:

1. No standardized tests that I know of can accurately measure qualities such as dedication, motivation, perseverance, caring, sensitivity or integrity.
2. We must admit the limits of tests and what they can measure. Moreover, we should recognize that tests must be limited in scope. They can present and measure only a sample of the knowledge required for teaching.
3. No test results guarantee that a prospective teacher will succeed and be a really good teacher in the classroom.
4. The rush to legislate excellence through teacher testing is raising some troubling signs and leading to some decisions that are educationally unsound.
5. One such decision, now law in several states, is to make continued accreditation of teacher preparation programs dependent upon the test performance of prospective teachers who are completing such programs. Educational Testing Service has testified against using teacher tests this way.
6. Some states allow only two years, from notice of probation, for graduating seniors to meet a predetermined standard of success on state certification tests. I believe this raises some of the same questions of fairness that have been raised in court challenges of testing programs for high school graduation.
7. A second area of concern regarding teacher testing has arisen in Arkansas and Texas. In the course of enacting comprehensive educational reform laws in both states, a requirement was included that all practicing teachers -- regardless of years of service and satisfactory ratings by their school supervisors -- would have to pass a one-time "functional academic

skills" or "literacy" test in order to retain their teaching certificate. Such a testing requirement is unprecedented for any other occupation requiring state licensure or certification.

8. To put an experienced teacher's professional career on the line solely on the basis of a mandatory, one-time test is both an injustice to the teacher and a misuse of tests. Educational Testing Service and the NTE Policy Council, in an unprecedented action for test development organizations, have refused to allow the use of NTE tests for this purpose in either Texas or Arkansas. (pp. 3-6)

The knowledge necessary for a teacher to function effectively in the classroom is not yet identified. The knowledge base needed by a kindergarten teacher and the appropriate teaching skills necessary to apply this knowledge to kindergarten cannot be measured by a pencil-paper test. The tests used today to make decisions about teacher competence are not designed to measure both knowledge and performance. The issue of validity in test development remains unsolved (Harnisch, 1985).

Many educators simply do not know much about testing and its application ("Testing Fears," 1985). By focusing on the results of tests, the major issue of proper use is sidestepped. The idea that testing is synonymous with "educational quality" is erroneous. Mandated testing becomes utilitarian in nature. The school becomes a factory developed to insure social efficiency. It has workers dedicated to the factory concept of satisfying its consumers. Students with passing scores insures teachers' jobs security and a guarantee against purchasing "faulty goods." But teachers are not workers on an assembly line and students are not equipped with warranties (Madaus, 1985). The knowledge and skills acquired through a good education cannot be measured by paper-and-pencil tests.

When testing is used in an accountability framework, it becomes basically punitive, demoralizing, and sends insidious and erroneous messages to people about ways of educating human beings. Evaluation is not synonymous with testing. Accountability calls for a higher order of ethical responsibility on the part of those who mandate. Demanding more time, more courses, more testing, more standards, and greater accountability in a mechanized fashion is foreign to educating America's youth (Sirotnik, 1985).

Test scoring is also problematic. For example, writing tests are generally scored holistically. Evaluation is related to the discreteness of a scale with so few score points that an adjustment on the part of two or three

readers pressed to complete hundreds of written examples in brief periods can drastically effect pass/fail rates. Phillips (1985) suggests solving this problem by grading writing passages with a method other than holistic scoring.

Educators have challenged the validity of pencil-and-paper tests because they do not measure on-the-job performance. Experts have agreed that there is no good measure of the effectiveness of teachers' on-the-job performance. The disparity of agreement raises serious questions about teacher competency and examinations for entry into teacher education programs, certification, and credentialing (Lines, 1985; Pigge & Reed, 1985).

Connecticut will be administering a test called the Connecticut Competency Examination for Prospective Teachers (CONNCEPT) in October of 1985. The attention being given to a meaningful and more appropriate validation study is reflected by the following statement:

Research in this area suggests that the development and validation of adequate criterion measures of professional performance present fundamental conceptual and measurement problems. To the extent that the criterion is questionable, any conclusions drawn from a criterion-related validity study must be questioned to at least the same degree.

An analogy to this line of reasoning is that government regulations traditionally are designed to protect the public from certain dangers rather than to guarantee specific outcomes. For example, in acquiring a drivers license, you must pass a written test on traffic law and a performance test on some basic driving skills. However, scores on these tests do not provide accurate predictions of future driving performance. Therefore, the test provides some critical information but is not by itself a sufficient indicator of a person's effectiveness in practice. (Pechione, 1984, p. 15)

In the Delaware study (Echternacht, 1983), the use of a passing range in scores and the realization that each test has a standard error of measurement is a very positive note.

The application of paper-and-pencil tests to make major decisions about the status of educators is related to other testing efforts. The historical base of mental testing is a well established and research based effort. All empirical data on predicting job performance have shown mental ability tests to be poor estimates (Callender, 1985).

Tests are limited in what they can measure; no test can

guarantee a good "anything." No standardized test can measure teacher qualities such as dedication, motivation, perseverance, caring and sensitivity. These qualities are necessary in a good teacher, according to Gregory Anrig ("Denver annual meeting," 1985).

Even though experts agree that there is no good measure of teaching effectiveness, the current trend to develop and implement more tests continues (Lines, 1985). Popham (1985) indicates that there is not an across-the-board effective teacher. He states:

The distressing truth regarding our teacher evaluation technology is that it is far more fragile than most people believe. We simply do not possess the requisite arsenal of measurement ploys needed to get a good fix on a particular teacher's instructional prowess. Without exception, there are problems in the data-gathering procedures we can employ to gauge a teacher's instructional skill. Those who think that we can comfortably rely on one or more meaningful measures of a specific teacher's merit are unaware of reality.

Irrespective of rationale, however, all of the recently installed teacher evaluations mandates are predicated on the same belief, namely, that we know how to evaluate teachers. That belief, in my view, is mistaken. I do not believe that a defensible technology for the appraisal of teachers currently exists. Moreover, I believe that the naive implementation of large-scale teacher evaluation systems may, in the long term, have an adverse effect on the quality of education.

Finally, because a good many large-scale teacher evaluation programs employ flawed procedures, a teacher's classroom procedures may be stultified because of the pressures of an appraisal system whose criteria are ill-conceived. (pp. 2-3)

#### The Impact of Testing on Minorities

Teacher education testing is having a devastating effect on ethnic minorities (Anrig, 1985; Brown, 1985; Ekstrom & Goertz, 1985; Goertz & Pitcher, 1985; Popham, 1984; "Teacher exams," 1985). The pass rates of ethnic minority groups on tests for entry into teacher education programs, credentialing, and certification continue to restrict the numbers of ethnic minorities entering and remaining in teaching (Goertz & Pitcher, 1985). Teacher shortages in certain types of inner-city school districts have already been noted. Teacher shortages will continue to be a problem for ethnic minority inclusion in the nation's teaching corps (Ekstrom & Goertz, 1985; Smith, in press).

A significant step was taken in California to insure greater access to Hispanics into the higher educational system. Chancellor W. Ann Reynolds of the California State University reports that although 25% of the state's school enrollment is Hispanic, only 10% of the university system's students are Hispanic ("Cal. Panel Urges", 1985). She indicates that, "California cannot afford this massive and growing loss of talent" (p. 2). Chancellor Reynolds appointed a Commission on Hispanic Underrepresentation, a panel of the 19 campus California State University system, in May 1984. The Commission has recommended 46 ways to increase the enrollment of Hispanic students. "Some of the recommendations involve major programs with substantial costs; others are low in cost and/or more narrow in scope" (Arciniega & Morey, 1985, p. 37). Three of the recommendations are as follows:

1. Programs to help high schools with large Hispanic enrollment to strengthen their college-preparatory courses and improve counseling.
2. The assignment of state university faculty members as "mentors" for Hispanic high-school students.
3. Programs to help Hispanics succeed on college entrance examinations. ("Cal. Panel Urges", 1985, p. 2)

According to Arciniega and Morey (1985), "The Commission has identified the following areas in which immediate action is necessary: (1) improved access to college, (2) direct services to students to assist them to progress educationally, and (3) the enhancement of institutional capabilities" (p. 37).

The poorer performance of minorities on competency tests for certification have been well documented (Lindahl & Wholeben, 1985). The mean scores of whites is 18 to 20 points higher than for blacks on all three parts of the NTE. Hispanics scored 0.6 to 0.8 standard deviations below the mean score for whites, and blacks scored 1.4 to 1.5 standard deviations below the mean for whites. The current trend will reduce the heterogeneity of the teachers in America (Anrig, 1985; Ekstrom & Goertz, 1985; Goertz & Pitcher, 1985). Passing rates for blacks and Hispanics have been as low as 10% to 50%. The pass rates of Anglos are consistently higher across the areas of math, reading and writing.

The Washington Post (Richburg, 1985, June 28) reported that 10% of the teachers failed the statewide competency test given to teachers in Arkansas. Out of 28,000 public school teachers, 2,803 failed the test. About seven percent

failed the writing portion, five percent failed the reading, and three percent failed the math. In the predominantly black Lee County, 34.5% of the teachers failed the test, while in Carroll County, with a mostly white population, 2.6% of the teachers failed. "When Louisiana decided to require teacher education candidates to pass the National Teacher Examinations, Grambling State U. found itself where many black colleges do--at the bottom. Fewer than 10% of its students qualified" ("Up From the Bottom," 1985, p. 238).

### Impact on Teacher Education Programs

Closing teacher education programs in institutions with open admission criteria which serve primarily minority populations will worsen the socio-economic and racial/ethnic mismatch between students and teachers. The current teaching profession is 87% white, 10% black and 2% Hispanic. States having large ethnic minority populations, such as California, New York, Texas and Arizona, have current testing programs which assure a decline of those minorities entering colleges with aspirations of becoming teachers (Goertz & Pitcher, 1985).

Ekstrom and Goertz (1985), Goertz and Pitcher (1985), Goertz, Ekstrom and Coley (1984), Smith (in press), and Kidd (1984) indicate that testing policies in most states have been mandated by legislators responding to the "excellence in education" movement. A review of actions taken by legislative bodies suggests that state policies are shortsighted and unawareness of the impact they will have on the teacher supply and on ethnic minority representation in American education (Ekstrom & Goertz, 1985). Testing continues to be the cheapest way to evaluate teachers but it is not likely to improve the caliber of those wishing to enter the profession. Instead, it simply denies access to those who score low on competency examinations ("Teacher Exams," 1985).

Peterson (1984) from ETS makes an important point which concerns many educators across the country who are dedicated to honesty in test application:

I consider much teacher testing policy to be badly conceived, hastily and/or mindlessly conceived, put forward for the wrong reasons, and then often finally shaped to the interests of the strongest power blocks. All of this, needless to say, is not unique to education policy-making. Someone once said there are two things you don't ever want to see being made: One is sausage, the other is law. (p. 15)

## Impact on Multi-lingual Populations

Hispanic-Americans make up the largest ethnolinguistic group in America. The disparity in the performance of Hispanics and whites on standardized achievement tests has long been known. The findings of Friedman (1985) confirm the fact that students tested in their dominant language score higher than when tested in their non-dominant language. The findings of other researchers verify language as a factor in state wide testing of ethnic populations.

Low social class is often associated with ethnic minorities. According to Campbell, Cunningham, Usdan, and Nystrand (1980):

The correlation between social class and academic success is well known. Students from middle- and upper-class backgrounds are more likely than lower-class students to do well on achievement tests, graduate from high school, and go on to college. On the other hand, lower-class students are more likely to enter school with severe educational deficiencies, behave in ways which middle-class teachers consider inappropriate and leave school at an early age. (p. 308)

Test scores for the CBEST in California revealed the same problem of higher failure rates for ethnic minorities. The pass/fail rates were similar in California on the CBEST as in Texas on the PPST, with the exception that the pass rates were even lower in Texas for minorities. This is explained by the different populations who took the test. In California, the early examinees included large numbers of teachers seeking certification, as compared with students seeking admission into teacher education programs in Texas. Students coming from educationally impoverished backgrounds, where English was not spoken, scored lower on the CBEST. In Georgia, students from historically black colleges scored sufficiently low to place teacher training programs at these institutions on probation. In California, teacher candidates scored higher than employed teachers which, perhaps, results from lack of familiarity with the material tested (Peterson, 1984).

Remediation for these tests becomes nearly impossible. The PPST, CBEST, and CAT are difficult to teach because institutions cannot specify deficiencies within each basic skills subtest. There is no evidence to indicate that institutions are coordinating their efforts to provide needed remediation (Ekstrom & Goertz, 1985). ETS also predicts only minor remediation possible based on its data base of PPST test results. Forty percent of the retakes on the PPST fail to change their scores (Goodison, 1985).

## Impact on Teaching Populations

States with large ethnic populations already underrepresented in the teaching profession, such as California and Texas, cannot afford to have such a non-diverse composition in school staffs. The problem is national in scope with continued implications for broadening the gap of participants in a society that is supposed to guarantee equal opportunity (Brown, 1985). Ethnic populations enrolled in institutions of higher education are underrepresented (see Appendix I for data).

The work of Brown (1985) and Lindahl and Wholeben (1985) supports the following statement:

Gregory Anrig, president of the Educational Testing Service, said testing can be used as one measure of qualification for teaching. But he and others worry tests might be used to determine competency of certified teachers. "It is just plain wrong to tell a teacher judged satisfactory for 10 or 15 years that the passing of one test on one day is necessary to keep his or her job or salary as a teacher," he said. ("Teacher Exams," 1985, p. 145)

Responses from organizations representing minority groups have continued to raise opposition to the use of those teacher competency tests which they believe discriminate against them. The voices of those denied entrance into the teaching profession could be transformed into rage (Popham, 1984). The existence of underprepared students who do not belong in college and the problems of requiring something different from minority groups are pointed out by Rabianski-Carriuolo (1985) and McCurdy (1985). Ethnic minorities have not had equal opportunities in this country (Brown, 1985). Failure rates and unfairness in testing are of great concern to many (Anrig, 1985; Raspberry, 1985).

Oakes (1985) and others propose that the concept of equity in America has been shelved in order to increase productivity, competitiveness, and individual "excellence." She proposes that the relationships between excellence and equity warrant further critical review. The very basis of our democratic belief in equality goes counter to the reversal of educational opportunity for ethnic minorities. This is an abomination.

### Cause for Failure

Explaining why ethnic minorities do not score as high as whites on tests involves a number of interrelated factors.

There seems to be a clear pattern established in the very early years of childhood. Children learn and interact in an environment which may be substantially different in language and custom from that in which they are supposed to function in order to succeed in life. The skills required to succeed in the White-Anglo English speaking world are often different than those apparent in the low socio-economic household with dissimilar communication skills, languages and customs. However, our tests contain language patterns and cultural orientations normed in an American society with Western European values and use English as the only language base. Lacking in use and practice, minorities face a greater challenge to interpret the written word in competition with others whose only spoken language has been English. Part of the solution may involve teaching minorities how to pass examinations. Testwiseness is a cognitive ability which can be taught and can be effective under certain circumstances (Dolly & Williams, 1985).

The complexity of cognitive activity related to scoring high on examinations can best be understood by the functional analysis necessary to decode written and spoken messages. The knowledge base acquired over an extended period of time becomes a memory bank for recall and interpretation of environmental cues which give our lives meaning. Cognitive processes and listening skills are enhanced with greater understanding as the levels of cognitive processing increase. The greater the depth of processing, the better the understanding and ability to interpret written and spoken information (Williams, 1985).

Cognitive entry skills of minorities are less refined and developed in a non-dominant language. Lack of practice and familiarity with values associated with English inhibit application of performance of routine operations, computation, reasoning skills and problem solving abilities. Part of the knowledge base in a different language and culture are not applicable and cannot become part of what is being measured in English. Most Hispanics, with cultural and language bases other than English, differ from Anglos in their respective profiles of cognitive skills. Recommendations should be on remediation, proactive compensatory instruction and the identification of relevant cognitive profile variables with well established predictive values of tests (Garcia, 1984; Lindahl, 1985).

Research bears out the most probable causes which prevent ethnic minorities from scoring as high as whites. The following conditions have been well documented: high ethnic minority school dropout rates, lack of parental interest, lack of English proficiency skills, reduced classroom participation, lower academic expectations,

unfamiliarity with test construction, test anxiety, inappropriate test-taking strategies, and lack of familiarity with test vocabulary. According to Reston (1985) Hispanic dropout rates are three times that of white and one and a half times that of blacks. Seventy percent of Hispanic children between the ages of 5 and 14 who live in Spanish-speaking homes have difficulty with English, and over 70% who remain in high school are enrolled in programs that do not prepare them for college. The known detrimental relationships among test performance, language background, parental socio-economic level, negative family and school experiences, and lower socio-economic status are additional factors which in varying degrees influence test performance. It is unknown how much each of the criteria influences test performance (Duran, 1983; Vasquez, 1985). However, the task of rethinking the testing mandate in this country may become a national agenda item when those calling for change understand how complex, inaccurate, and probably inappropriate, single test scores are in making decisions about individuals' professional lives.

#### Hispanic Population Data

The inability to attract talented minority students into the teaching profession continues to be a great concern. Minorities constitute 17% of the total population and 27% of the school age population. Stringent selection will continue to deplete an already underrepresented population. In California, 44% of K-12 grade children are from minority backgrounds ("Board Adopts NCETE Report," 1985; "Denver Annual Meeting," 1985). With the implementation of testing, it is anticipated that the minority teaching force will be cut almost in half by the year 2000 ("Denver Annual Meeting," 1985).

There is a contradiction between the "excellence" reports and the real nature of students in school. Blacks and Hispanics do not have equal access to the occupational structure of this country. Minorities are not permitted to receive the economic benefits from education (Weis, in press). The representativeness of Mexican-Americans in New Mexico improved slightly over the past decade but lost ground in Texas, with California roughly staying the same (Payan, Peterson, & Castille, 1984).

There are more than 15 million Hispanics in the U.S. with an escalation in immigration from Central and South America. Seventy-five percent live in five states: Texas, California, New York, Florida, and New Mexico. Four-fifths of all Hispanics live in households where Spanish is usually or sometimes spoken. One-fifth of all Hispanic families had incomes below the poverty level. Hispanics are

underenrolled in higher education and very seriously underrepresented, more seriously than blacks, in the teaching profession. Eighty-seven percent of all teachers are white, 10% are black, and 2% are Hispanic (Brown, Rosen, Hill, & Olivas, 1980). Compare this to Popham and Yalow's data on the ethnic distribution of students in Texas in 1983:

. . . there were 52.2 percent White students, 29.7 percent Hispanic students, and 16.3 percent Black students. Statewide, for the 1981-82 school year, the ethnic composition of Texas public school students was 55.7 percent White, 28.1 percent Hispanic, and 14.8 percent Black. (pp. 7-8)

Additionally, "the number of Hispanics in California will nearly double over the next 15 years . . . By the year 2010, a majority of the state's population is expected to be composed of Hispanics, blacks, Asians, and members of other minority groups (McCurdy, 1985a, p. 2)

According to Church (1985):

. . . the growth of the U.S. Hispanic population is one of the most startling phenomena in American social history, and if anything it is likely to speed up. . . . Last year there were an estimated 17.6 million, with roughly 60% tracing their ancestry to Mexico and the rest to Puerto Rico, Cuba, El Salvador, the Dominican Republic, Colombia, Venezuela, and about two dozen other countries of Central and South America.

Shortly after World War II, three-quarters of all Hispanics on the U.S. mainland lived in Texas or California. As of 1980, those two states still accounted for 51% of the total Hispanic population. But large numbers have also settled in Arizona (16% Hispanic) and New Mexico (36%) and in such inland and Northern cities as Denver (19%) and Hartford, Conn. (20%). In South Florida, nearly a million Hispanics (78% Cuban) have spread so rapidly beyond Miami (64% Hispanic) . . .

Some analysts think that Hispanic Americans by the year 2000 will total 30 million to 35 million, or 11% to 12% of all U.S. residents . . . (p. 36)

According to Smith (in press):

The National Center for Educational Statistics (as reported in Andrews, 1983) indicates that minority public school enrollment presently exceeds 50% in New Mexico and Mississippi. Minority public school enrollment is projected to approach 50% in California,

Louisiana, South Carolina, and Texas by 1990. By the year 2000, Alabama, Arizona, Georgia, Florida, and North Carolina, which currently have minority public school enrollments ranging from 30 to 40%, are likely to approach the 50% mark. (pp. 38-39)

### Legal Aspects

Freeman, Hess, and Kasik (1985) have discussed extensively the legal aspects of teacher testing. The courts have examined two well established legal principles which have impacted on testing across the nation. The first is the right of individuals to engage in the gainful occupations of their choice as protected by due process. The second is the state's authority to reasonably regulate in order to protect the public interest.

Griggs v. Duke Power Company (1971) reversed the "common sense standard" used up to that time which relied upon the credibility of testimony or on minimum attempts at validation (Freeman et al., 1985). In Griggs, the "court eventually decided that any test is illegal unless it clearly measures the skills needed for the job" (p. 3). Hunnicut v. Burge (1973) concluded that qualified teachers should possess "a good education." Freeman and his colleagues (1985) have discussed the difficulties in specifying legally what this might entail.

The decision in Washington v. Davis (1976) modified interpretations given by the courts earlier because intent rather than impact was stressed (Freeman et al., (1985). The court ruled that tests' job relatedness need not be considered unless actual intent to discriminate was shown. This has provided much greater latitude in test interpretation. When a test seems "obviously" relevant and was developed "in good faith," it is accepted and the Equal Employment Opportunity Commission (EEOC) guidelines for job relatedness are not considered (Freeman et al., 1985).

The U.S. v. South Carolina (1977, 1978) decision permitted the use of tests to certify minimally qualified persons through the use of a content validated standardized test (Freeman et al., (1985). The court ruled that "standardized tests do reflect individual achievement with respect to specific subject matter content which is directly relevant to (although not sufficient in itself to assure) competence to teach" (cited in Freeman et al., 1985, pp. 25-26). Hazard, Freeman, Eisendorfer, and Tractenberg (1977) stated that a general education is what each teacher should have, giving greater generalizability to test developers (cited in Freeman et al., 1985). Stewart v. Hannon (1979) dealt with licensing exams for principals and

upheld the use of examinations for this purpose (Freeman et al., 1985).

The out-of-court settlement of Golden Rule Insurance Company v. Mathias (1980) in November, 1984, is likely to have impact on future testing because the settlement, in effect, requires greater supervision of tests and their uses (Freeman et al., 1985):

This settlement which evolved subsequent to the Illinois Supreme Court's denial of defendants' petition to file appeals contains these fundamental points:

- (1) Applicants taking subsequent examinations will be asked to furnish voluntarily their race or ethnicity and the level of education they have attained.
- (2) The Director of the Department of Insurance will publish annually a report containing statistical information, including the results of the testing by race--results for each part and the entire test, the mean scaled scores on each part, and the standard deviation of scaled scores on each part, and the results by race for those having a high school diploma or a G.E.D. In addition, the Department is required to prepare an "item report" which provides the correct-answer rates for each item by race as well as r-biserial correlations.
- (3) In constructing future tests, ETS is obliged to adhere to both the APA's Standards for Educational and Psychological Tests and ETS' Standards for Quality and Fairness and Test Sensitivity Guidelines.
- (4) The agreement stipulates that no more than a twelfth grade reading level as determined by generally accepted reading-factor indexes, e.g., FOG or SMOG, shall be required (excluding terms specifically related to the insurance industry, e.g., [the term] beneficiary).
- (5) The construction of future tests, according to the agreement, is to proceed in the following manner:

All test items are to be divided into two types. Type I items are those for which (a) the correct-answer rates of black examinees, white examinees, and all examinees are not lower than forty percent (40%) at the .05 level of statistical significance and (b) the correct-answer rate of black and white examinees differs by no more than fifteen percentage points at the .05 level of statistical

significance. Type II items are all other items.

The agreement then stipulates that Type I items are to be used exclusively as long as there is a sufficient number. When used, Type II items are to be selected in descending order of the least differences between White and Black examinees. New items may be generated but must be pre-tested on three occasions; these items, known as Type III, cannot be used until after being pre-tested and then classified as Type I or Type II items.

- (6) Finally, the agreement stipulates the creation of an advisory committee comprised of representatives of the insurance industry in Illinois and two persons knowledgeable in the area of psychometrics. The committee will review test results and make recommendations concerning the test (pp. 11-12).

Some independent school districts have developed their own tests and validation studies. The Houston Public Schools developed the examination called the "Functional Academic Skills Test" (FAST) to assess basic reading, writing, and mathematics skills. According to W. N. Kirby, Commissioner of Education in Texas, of the 12,000 to 13,000 professional staff from the Houston Independent School District, 7,000 have passed the reading and writing parts of the FAST examination (Kirby, 1985). The "face validity" approval given to the FAST by the Texas Education Agency staff appears to be inconsistent with the guidelines of the Standards for Educational and Psychological Testing (1985).

The use of any test for employment selection, promotion, or retention requires that a test be appropriately validated by the agency administering the examination. Beyond this, there are professional, ethical, and legal responsibilities that the examiners must safeguard against should the examination have an adverse impact on hiring, promotion, or membership opportunities of members of any race, ethnic, or gender group. Employees may file suit against an employer, claiming discrimination in employment practice, if criteria and procedures for evaluation are not consistent with the Uniform Guidelines in Employment Selection Procedures, published in 1978 (Stein & Frankiewicz, 1985).

To be legal, a testing program must be equitable and fair. Examinations must measure what teachers should know to be effective teachers. Tests cannot be used to discriminate by race or gender. The safest possible legal ground for public administrators are indicated by Lines (1985) as follows:

1. Appropriate validation of tests.
2. Insure that tests measure what they are intended to measure and that this reflects skills needed for on-the-job performance.
3. Use tests only for purposes and applications recommended by the developer.
4. Avoid tests which disproportionately exclude a racial or gender group.
5. Use additional non-test appropriate criteria.
6. Insure adequate procedures for challenging errors or abuses.

It is important to recognize the differences between present testing criteria and legal criteria used by the courts for dismissal purposes over the past 40 years. The courts have accepted pre-employment tests with some reservations, and require that they be equitable and fair. Seventeen states have adopted a testing requirement for new teachers. Fourteen are considering general competency tests. The courts reportedly are willing to question efforts which are not a reasonable measure of performance (Lines, 1985). Texas has approved a state-wide test of basic skills to be taken by all teachers in the area of reading and writing (Langford, 1985).

The development and applicability of tests in South Carolina are technically (and thus probably also legally) defensible based on the Angoff, Nedelsky, and Ebel approaches (Hamm & Winter, 1985). These approaches require groups of experts to make item-by-item judgments of each test item difficulty. However, the process of asking experts to make judgments about actual performance as a method of measuring actual performance is not sufficiently accurate (Garcia, 1971).

The responses to and the reasons for mandating change in the testing movement in this country are developing:

1. In Florida, both teacher unions have filed litigation against the Master Teacher Plan. At least 10 other states have legislation pending or have passed similar legislation (G. W. Wilson, 1985).
2. There is insufficient research evidence favoring one method of job analysis for construction and validation of personnel tests indicating that there should be multiple methods used.
3. Tests developed on the basis of systematic job analysis are more fair than commercially developed tests.
4. Specific written knowledge tests developed through job analysis are more fair to minorities (Stein & Frankiewicz, 1985).

# BEST COPY AVAILABLE

## Solutions

The impact that teacher reform will have on the teaching profession is yet to be felt. The processes of test development and implementation have expanded beyond local and state levels. The new movement toward the development and implementation of "a national teacher examination" by 1990 has surfaced as being a possible solution to the testing dilemma.

The problem of establishing cut-off scores for "merit" systems is accompanied by "quota" systems with rewards from a fixed pool of resources. The criterion for making decisions must be clearly defined and observable in making on-the-job performance judgments. Cut-off scores as a sole criterion, rewards from fixed amount of resources, and criteria unrelated to performance should be declared void (Capie, Ellett, & Cronin, 1985).

Predictive validity which is related to performance in testing refers to the ability of an examination to predict performance which will occur at some later date. For example, McCaleb (1984) found that testing of specific oral communication skills satisfactorily predicted student teachers' oral performance. However, most of the research which will reveal empirical evidence on skills associated with teaching competence is yet to be done (McCaleb, 1984).

A performance profile could be created as a part of periodic evaluation of personnel. This process involves no cut-off scores with all information viewed as one criterion for evaluation (Cronin & Capie, 1985). The documentation of job-relevance beyond a credentialing examination is necessary. There are clear differences between instructional, curricular, and performance validities. Content validity based on materials and what is actually taught to students is quite different from performance validity which deals with job-related behaviors (D'Costa, 1985).

Evaluations of teachers should be made at specific periods during their preparation periods. Different examinations covering different criterion domains should be used to insure that correct assessment of learned information performance takes place. Medley's (1985) work on developing a competency-based teacher certification system for the state of Virginia brings important clarity to types of evaluation problems associated with validation of teaching and teacher professional behaviors. He identifies five major targets (periods in a teacher's career) when testing is most appropriate because different domains would be assessed:

1. Teacher evaluation may focus on personal characteristics of candidates for admission into pre-service professional teacher preparation programs.
2. Teacher evaluation may focus on knowledges, skills, attitudes, etc., that teachers possess at the end of a period of professional training.
3. Teacher evaluation may focus on the actual behaviors of teachers while teaching.
4. It may focus on the behavior of a teacher's students while they are being taught.
5. Finally, teacher evaluation might focus on the knowledge, skills, attitudes, etc. that the teacher's students possess after the teaching stops.

I will refer to these five foci as "assessment points" and number them one to five, respectively.

Pre-existing teacher characteristics are assessed directly at Point 1. They include characteristics of two kinds, first, knowledge, skills, attitudes, etc., that the candidate will need in order to satisfactory work in the teacher preparation program; and, second, knowledge, skills, attitudes, etc. the candidate will need in order to succeed as a teacher, but will not acquire as part of her professional training.

Teacher competence is assessed directly at Point 2. A competent teacher is one who possesses a minimum set of teacher competencies. A teacher competency is a unit of knowledge, a skill, or a set of attitudes required for satisfactory performance as a teacher.

Teacher performance is assessed directly at Point 3. Teacher performance refers to the pattern of behavior a teacher displays while teaching a class.

Teacher competence is a personal characteristic of the teacher that the teacher carries with her from one setting to another, but teacher performance is a function of setting and will change from one setting to another.

Student learning experiences are assessed directly at Point 4. Student learning experiences are activities of students that result in learning.

Student learning outcomes are assessed directly at Point 5. Student learning outcomes are changes in student knowledge, skills, attitudes, etc. measured after teaching ends. The related term teacher effectiveness refers to the portion of these outcomes that is attributable to the efforts of teacher. (pp. 1-3)

The process of implementing an evaluation program to assess classroom performance often involves political, financial, and educational objectives that are inconsistent and incompatible with each other. Drawing on a multitude of

research, Garrett (1985) points out that:

Ideally, an acceptable evaluation system should be selected and supported by the teachers involved, encourage teacher and principal participation in establishing performance goals, provide opportunity for the feedback of information to teachers, make provision for a concluding reappraisal of the system by the participants, focus on the processes of teacher development, be known to teachers in advance of its application, and, most importantly, espouse a philosophic base which is directly linked with the improvement of instruction. (pp. 1-2)

According to Hunter and Slaughter (1980), "Educational Testing Service is committed to the development of tests that reflect a thoughtful and humanistic consideration of all people and acknowledge the multicultural nature of our society" (p. 5). It is also noted that vigilant efforts must be made to evaluate ETS tests from the perspectives of Asian Americans, Black Americans, Hispanic Americans, Native Americans, and women. A claim of test item analysis on the basis of cognitive/affective nature is made which would seem impossible to confirm. Tests are being redesigned by an assortment of groups in many states with ETS and others often playing a limited role.

The writer reviewed two forms of the PPST and one form of the CBEST and found the tests to be free of terms which would reflect negatively on the minorities mentioned above. However, no claim is made by ETS that the tests had been reviewed for language and cultural bias by these groups. The opportunity was afforded the writer to identify colleagues with expertise in Spanish to review the PPST for language and cultural bias. To the writer's recall, there were no questions related to the language or cultural heritage of the American Hispanic population on any of the tests reviewed. Most references made in questions related to our American-Western European cultural heritage.

Suhor (1985) reports that teachers have more confidence in observations and teacher-made tests than in standardized tests. There is a small amount of empirical evidence which supports the testimony of practitioners. The conclusion is made that reliance on objective tests is caused by inertia, ignorance, and lack of funds. The cost of more complex evaluation and the possibility of teachers legally challenging less "objective" measures may also be factors.

There is considerable confusion about what should change at Institutions of Higher Education (IHEs). Essentially, changes have come about in teacher education programs where

mandates have been made. In some states, such as California, only minor changes have been made. The variation of state policies raises some ethical questions about the diversity of applicable criteria deemed necessary in the rush to mandate teacher competency (Carlson, 1985). There are innumerable combinations of criteria or standards that can be applied and can challenge a sense of fairness in evaluation. To complicate the testing picture further, it should be noted that filters are applied to different groups of students entering college at different levels. Different criteria are used to evaluate entrance with different cut-off scores and applied with at different levels of the teacher training programs. Testing programs vary considerably with some being custom designed for specific populations and others using available standardized tests. Caution in test use is a must (Coley, 1985).

Schlechty (1985) proposes a framework for inducting or developing in new members of the teaching profession the skills, knowledge, attitudes, and values necessary to carry out their occupational goals. His work suggests that all occupations have prescribed occupational goals. Evaluation through "induction into the teaching profession" by the demonstration of accepted knowledge and the performance of certain skills could bring order to a complex and varied testing environment.

#### The Future

Evangelauf (1985) reports that the Southern Regional Education Board (SREB) has developed a new "covenant" as a solution to solve the problem of the unacceptably low quality of undergraduate education. The findings cited as justification for the SREB recommendations indicate that 40% of college freshman needed remedial work and that the average community college freshman is reading at the eighth grade level.

The actions recommended by the SREB included state and university controls on quality programs which support remedial programs provided for unprepared students which in effect would insure access (Evangelauf, 1985).

The political, fiscal, and "excellence" motives molding the future of education hold both promise and disappointments. As the restructuring takes place, it has become apparent that various elements interacting with each other have the potential for changing education in numerous ways. Testing is one of those elements that cannot be viewed as an isolated force. For example, in South Carolina, the Education Entrance Examination (EEE) was implemented without the realization that large numbers of

students would fail, seriously affecting the existence and structure of education programs (Hamm & Winter, 1985).

The future of education will be influenced heavily by national groups attempting to steer the movement toward concrete positions. We are entering the shakedown phase where organizations such as the American Federation of Teachers (AFT), National Education Association (NEA), American Association of Colleges of Teacher Education (AACTE), National Council for the Accreditation of Teacher Education (NCATE), and more recently the Holmes Group Consortium are setting their agendas for joint developments in testing and teacher training (Jacobson, 1985c, 1985d).

The need for teachers to take greater control of their profession is an important element in the response to public concern for teacher competence. Testing teachers and removal of incompetent educators have become the byword for union leaders in a search for a structure under which education can flourish with self-respect (Currence, 1984a).

According to Feistritz (cited in Currence, 1984b) teacher certification procedures are "a mess" and should be replaced by national certification standards. The requirements for different types of certification vary from state to state. Teacher education programs also vary from state to state. Licensure in one state permits a teacher to teach certain grades in one state but not in another state. Feistritz's survey revealed:

- . Forty-eight states and the District of Columbia include graduation from an "approved program" among their eligibility requirements for certification.
- . Although all states require a baccalaureate degree, the credit-hours demanded in general studies, professional-education courses, and clinical experience vary widely from state to state.
- . Certification from state to state lasts anywhere from one year to "life."
- . All but two states, Vermont and Virginia, issue substandard, limited, or emergency credentials to people who do not meet all of the criteria for certification. Half of those states will give a substandard credential to people who have less than a bachelor's degree.
- . All but 18 states already are in some stage of examining alternatives to the traditional teacher-training program route to certification, and Florida and New Jersey are considering proposals for certifying teachers that would completely circumvent teacher-education training [as of this writing, New Jersey has enacted this legislation]. (pp. 1, 10)

The NEA still opposes the use of test scores as sole criterion for teacher certification but recently softened its anti-testing position. It is likely that new directions in teacher certification will require multiple evaluation methods ("Special Panel Suggests," 1985). The establishment of five year teacher education programs has gained broad support at the national level. More emphasis is also being placed on a broader liberal arts education in teacher preparation. The changing emphasis on traditional teacher education programs would supposedly force screening processes for applicants into teacher education programs (Jacobson, 1985a; 1985b; "Special panel suggests," 1985).

The career ladder and merit pay have been suggested to reward superior teachers. Tennessee appears to be the leader in successfully implementing a comprehensive career ladder program. Multiple criteria will be used in the program's evaluation of teaching effectiveness. Even though programs such as these reflect greater professional credibility, the shortcomings of testing and evaluating teaching performance remain problems of great concern. There are numerous areas, domains, and indicators which are not observable in a teaching-learning situation (Furtwengler, 1985).

Corrigan (1985) states that the career ladder is actually restructuring the very nature of teacher education. Colleges of education cannot go out on their own in designing programs that are unrelated to career ladder systems that determine salaries, certification, and training needed to succeed in them.

The Secretary of Education has lauded recent attempts to establish a national test which will restore professionalism to teaching ("Bennett," 1985). The proposed test would be somewhat different than the National Teachers Examination, with greater rigor than existing tests at a cost of about \$200. The new test would consist of written essays, multiple choice questions, a performance component, and require several days to complete (Imig, 1985).

PART 2

NATIONAL SURVEY ON TESTING

Background

In the spring of 1985, a questionnaire was sent to the director of teacher certification and to a selected dean of education in each state requesting information on state-wide use and impact of five categories or target areas of teacher testing--(1) admission to teacher education; (2) professional education; (3) academic; (4) on-the-job; and (5) certification. Two questionnaires were sent to each state to insure greater accuracy in data collection and to attain a higher return rate. Responses were secured from all 50 states (see Appendix J for a sample questionnaire).

Overall, the data confirmed the trends reported by Sandefur (1985) in his survey of these same groups. It should be pointed out that this study and Sandefur's collected complimentary data, often from identical informants. Thus, in many aspects, the results of the two studies complement one another.

Several state informants indicated that they were at stages near test planning or implementation. Some had identified specific future dates when testing would take place. This data was consistent with that data gathered by Sandefur (1985). He has analyzed the number of states presently mandating and planning competency tests. His tables appear in Appendixes A and B.

Targeted Areas

Results on only those testing programs to be actually implemented by September 1985 are reported here. The number of states reporting testing in each category were as follows:

- 1. Admission to Teacher Education . . . . 23
- 2. Professional Education (Pedagogy). . . 14
- 3. Academic . . . . . 16
- 4. On-the-Job . . . . . 8
- 5. Certification. . . . . 22

The greatest number of states reported using tests for admissions (23 states), followed by certification (22 states), academic (16 states), pedagogical testing (14 states), and "on-the-job" (8 states). Those listed in the "on-the-job" category included some such as "career ladder observations" which technically may not be classified as identifiable state mandated examinations.



## Perceptions of Test Content

For each category or target area of teacher testing, respondents were asked to list the names of tests used within this area and then to respond to the question "What does the examination test?" by selecting one or more of seven categories.

Each state was surveyed on perceived test content in order to identify the intended use for each test's reported use. Test developers such as ETS provide literature recommending appropriate use of their tests. However, it is the states who have legal responsibility for test validation and appropriate use. The perception of what the tests measured and the test developer's intended uses were sometimes different. The diverse responses reported for a given test raise some questions about how tests may be used.

Table 2 contains the number of responses to the question "What does the examination test?" indicated for each targeted category area. In many instances, a given test was reported to be measuring two or more areas. Some states indicated a large portion of targets being assessed with the same test. In some instances, respondents identified single tests as measuring several of the seven content areas. It is possible that educators at the highest levels may, in fact, know little about what those mandated tests actually measure.

Table 2

States' Responses to "What Does the Examination Test?"

What tests are testing (Respondents)	Categories of Testing				
	Admission into Teacher Education	Professional Education (Pedagogy)	Academic	On-the-Job Performance	Certification
Basic skills	26	7	1	1	11
If a student can perform in a teacher education program	4	1	-	-	2
Knowledge a student should have to do well as a teacher	5	1	10	1	11
Teacher competence-- knowledge and a set of attitudes required for satisfactory performance as a teacher	1	11	3	5	10
Teacher performance-- a pattern of behavior a teacher displays while teaching a class	-	1	-	8	1
Teacher performance based on student learning outcomes-- changes in student knowledge	-	-	-	1	1
Other	3	6	7	-	4

## Test Use as Reported by States

The questionnaires received from state certification officers and heads of education units reported a diversity of uses for identical or similar examinations. Some examinations were used with multiple populations in some states and with entirely different populations in other states. For example, the PPST is used by Texas for entry into teacher education programs and by Delaware for teacher certification.

Table 3 indicates the specific tests named by the respondents and the number of states which selected each category of content use. Those surveyed were asked to respond to the question, "What does the [specified] examination test?" The content categories listed below were provided for each respondent to decide what area(s) their examinations tested. The number of states which indicated each category is listed on the table below each category letter.

- A. Basic skills
- B. If a student can perform in a teacher education program
- C. Knowledge a student should have to do well as a teacher
- D. Teacher competence--knowledge and a set of attitudes required for satisfactory performance as a teacher
- E. Teacher performance--a pattern of behavior a teacher displays while teaching a class
- F. Teacher performance based on student learning outcomes--changes in student knowledge
- G. Other (specify)

Table 3

States' Responses to Content/Use Categories for Each Test Listed for Target Areas of Competency

Test Name	Number of States Using Test	<u>Content/Use Categories</u>						
		A	B	C	D	E	F	G
Alabama English Language Proficiency Test	1	1		1				
ACT	3	3		3				1
SAT	3	3		1				1
PPST	9	9	2	1				
Arizona Teacher Proficiency Examination	1	1						
CBEST	2	2						
CAT	1	1						
Connecticut Comprehensive Examination for Prospective Teachers	1	1		1	1			
Florida's College Level Academic Skills Test	1	1						
CTBS	1	1						
NTE	3	2	1	1				
ACT College Outcomes Measurement Program	2	2	1	1				
State Mandated Basic	1	1						
Education Entrance Examination	1	1						
Washington Pre-College Test	1	1						

Table 3 (cont.)

<u>Professional Education (Pedagogy)</u>								
Test Name	Number of States Using Test	<u>Content/Use Categories</u>						
		A	B	C	D	E	F	G
Basic Professional Studies Test of the Alabama Initial Teacher Certification Testing Program	1	1			1			
Arizona Teacher Proficiency Examination	1				1			
NTE	10	5	1		8		1	6
Florida Teacher Certification Examination	1	1			1			1
Assessment of Performance in Teaching	1							1
South Carolina Teaching Area Examination	1							1
Career Ladder Professional Skills Test	1	1	1		1			

Table 3 (cont.)

Test Name	<u>Academic</u>							
	Number of States Using Test	<u>Content/Use Categories</u>						
		A	B	C	D	E	F	G
Alabama Initial Teacher Certification Testing Program	1				1			
Individual by subject area	1			1				
NTE	9	1		6	2			5
Teacher Certification Test	1			1				1
Curriculum examinations	1			1				1
South Carolina Teaching Area Examinations	1							1

Table 3 (cont.)

Test Name	<u>On-the-Job</u>							
	Number of States Using Test	<u>Content/Use Categories</u>						
		A	B	C	D	E	F	G
Arizona Educational Skills Assessment Test	1	1						
Florida Performance Measurement System	1				1	1		
Teacher Performance Assessment Instrument	1		1	1	1			
Georgia Performance Appraisal	1			1	1			
Entry Year Teacher Observation Instrument	1					1		
Assessment of Performance in Teaching (APT)	1					1		
Career Ladder Evaluation System	1			1	1			

Table 3 (cont.)

Test Name	<u>Certification</u>							
	Number of States Using Test	<u>Content/Use Categories</u>						
		A	B	C	D	E	F	G
Basic Professional Studies Test of the Alabama Initial Teacher Certification Testing Program	1	1			1			
Arizona Teacher Proficiency Examination	1	1	1	1				
NTE	16	6		8	9	1	1	11
CBEST	2	2						
CAT	1	1						
PPST	2	2						
Florida Performance Measurement System	1						1	
Georgia Teacher Certification Test	1			1				1
Oklahoma Teacher Certification Exams	1			1				1
South Carolina Teaching Area Exams	1							1
ACT College Outcomes Measurement Program	1	1						

## Impact on Minority Groups

For each test listed, respondents were asked to report known pass/fail rates for specific ethnic groups ("White", "Black", "Hispanic", and "Other") and to compare the passing percentage rate of whites with those of "minorities". Eleven states provided pass/fail data categorized as follows:

- Part 1. Admission into Teacher Education
- Part 2. Professional Education (Pedagogy)
- Part 3. Academic
- Part 4. On-the-Job Performance
- Part 5. Certification

Table 4 summarizes the pass/fail data for those eleven states who included the ethnic data.

Table 4

### Pass/Failure Rates by Ethnic Group

State/ Part	White		Black		Hispanic		Other	
	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail
AZ 1	80	20	44	56	56	44	70	30*
AZ 2	99	1	91	9	96	4	68	31*
CA 1	76	24	30	70	38	62	-	-
CA 5	76	24	30	70	38	62	-	-
CO 1	98	2	95	5	97	3	-	-
FL 2	90	10	40	60	55	4	-	-
FL 4	95	5	90	10	-	-	-	-
GA 5	94	6	54	46	-	-	-	-
MS 1	70	30	40	60	-	-	-	-
NC 2	97	3	72	28	-	-	-	-
NC 3	97	3	72	28	-	-	-	-
NC 5	97	3	72	28	-	-	-	-
OK 3	78	28	45	55	71	29	70	30**
OR 1***	70	30	-	-	10	90	-	-
SC 4	100	0	100	0	-	-	-	-
NM 1	58.8	41.2	42.2	57.8	42.2	57.8	42.2	57.8**
NM 2								
Comm Sk	97.7	2.3	50	50	92.5	7.5	60	40**
Gen Kn	97.6	2.4	80	20	88.2	11.8	50	50**
Prof Kn	- Insufficient sample of numbers of students passing the test							

\*Other - Asian Americans

\*\*Other - Native American

\*\*\*First attempt to obtain pass/fail rates, small sample

A comparison of the pass/fail rates indicates a lower pass rate for ethnic minorities in each state reported. This confirms earlier data reported by Brown in 1985 (see Appendix K).

Most states not reporting pass/fail rates gave the following reasons:

1. Do not have the information
2. Not applicable
3. Not available
4. The state department has that information
5. Do not gather this data

Sixteen states responded to the less specific question asking only for a comparison of "White" and "Minority" pass rates. Table 5 contains the data received.

Table 5

A Comparison of White vs. Ethnic Group Pass Rates

Test Category	<u>Number of States Responding to Request</u>		
	Lower	Equal	Higher
Part 1 - Admission into Teacher Education	16	-	-
Part 2 - Professional Education (Pedagogy)	5	-	-
Part 3 - Academic	7	-	-
Part 4 - On-the-Job	-	-	-
Part 5 - Certification	7	-	-

All 16 states indicated that ethnic minorities scored lower on teacher education admission tests. Five states reported lower test scores for ethnic minorities on tests of professional education (pedagogy). Seven states reported lower test scores for minorities in the area of academics. Seven states reported lower test scores on certification tests. The only area not reporting any results of testing was in the area of on-the-job performance. Not a single state reported that minority groups scored equal to or higher than whites in any category.

## Questions of Multiple Criteria

The AASCU study on entry requirements into teacher education (Sikula, 1984) indicated that multiple criteria have continually been used for entry into teacher education programs. Written responses to the questionnaire in this study suggested that states were continuing to use multiple criteria. However, the follow-up telephone survey conducted to clarify written responses suggested something quite different. Those states whose respondents had reported the use of cut-off scores were called. Although multiple criteria were, indeed, being used for admission into teacher education, students scoring below the established cut-off test scores were excluded because of the single criteria. Thus, the introduction of testing with specified cut-off scores has adversely affected educational opportunities for ethnic minorities and has, in fact, served as a sole criterion for denial to programs.

## PART 3

### AN ANALYSIS OF THE PPST AS USED IN TEXAS

#### Introduction

In 1979, the 66th Texas Legislature created the Commission on Standards for the Teaching Profession, a representative group of professionals charged with responsibility for studying teacher preparation and certification. As a result of recommendations by the Commission on Standards and the State Board of Education, the 67th Legislature in 1981 passed Senate Bill 50 which mandated the competency testing of students prior to entering teacher education programs and the competency testing of students completing teacher education programs prior to certification as teachers, administrators, or professional educators.

In 1982, the Texas State Board of Education adopted the Pre-Professional Skills Test (PPST) published by Educational Testing Service (ETS) as a basic skills screening device for admission into Texas teacher education programs. In January, 1983, the Commissioner ordered a study to evaluate the content of the PPST. Between February 2, 1983, and September 6, 1983, six validation reports on the PPST were submitted to the Texas Education Agency (TEA) by IOX Assessment Associates of California.

IOX Assessment Associates' data on the PPST determined its appropriateness as a screening device for entry into teacher education programs in Texas. According to Popham and Yalow (1983) of IOX Assessment Associates:

Prior to its use in Texas, the P-PST had not been widely employed. In fact, much of the data used to confirm the psychometric quality of the P-PST was acquired during the TEA-commissioned investigation of the examination. In part, it was the recency of its development that led TEA officials to gather evidence of the P-PST's quality. More fundamentally, however, it was recognized that if the P-PST were actually to be used to exclude individuals from Texas teacher education programs, the possibility existed that the examination's adequacy would be challenged in court. TEA officials wished to assemble evidence regarding the P-PST's quality so that, were the examination technically defensible and appropriate for use in Texas, its use could withstand legal scrutiny. If, on the basis of the P-PST study, it appeared that the examination's quality was questionable, then educational policymakers in Texas

would need to consider alternate examinations or to re-think the 1982 screening-examination policy. (pp. 2-3)

### Determining PPST Suitability

The three major categories used to study the suitability of the PPST were:

1. Adequacy-of-preparation
2. Content validity in terms of "program" and "job" relevance
3. Selecting performance standards

#### Adequacy-of-Preparation

Adequacy-of-preparation refers to the extent to which examinees have been adequately prepared to successfully pass the PPST. Adequacy-of-preparation becomes important because it is unconstitutional to test, in order to determine access, on information not taught to prospective examinees. According to Popham and Yalow (1983):

Previous court rulings regarding high school graduation tests (e.g., Debra P. v. Turlington, 1981; 1983) and the use of the National Teacher Examinations (NTE) as screening examinations (e.g., United States v. South Carolina, 1977) have mandated that it is unconstitutional for an agency to deny access to a legitimate expectation (for example, a teacher credential) because of a test if that agency had responsibility for, yet failed to teach, the contents of the test. Accordingly, the initial task of the investigation was to assemble evidence regarding the extent to which preparation for the P-PST had been provided. (p. 4)

Three sources of data were used to provide evidence about the extent to which individuals had been provided the necessary instruction to pass the test--Texas public school educators, Texas college students and textbooks used in Texas secondary schools. Texas public school educators were asked to make judgments about the content of the PPST as follows:

Panelists reviewed each section of the P-PST both on an item-by-item and a total-test (total-section) basis. For the item-by-item judgments, educator panelists were required to answer the following question for each multiple-choice test item in the section(s) of the P-PST that they reviewed: "Do students in the college-preparatory program in your district have the

opportunity to learn the content of this item?" Panelists were to respond by circling Yes, No, or an asterisk (indicating insufficient information to make this judgment) for each test item. For the total-test judgments, educator panelists were asked to select one of four statements "almost all," "most," "a small amount," or "almost none" that best reflected their judgment regarding the proportion of the section(s) of the P-PST that they reviewed for which preparation was provided in their district's college-preparatory program. (Popham & Yalow, 1983, p. 8)

Panelists were given the following instructions:

Do not try to judge whether students, in fact, know the information; only judge whether they have the opportunity to learn it. Please keep in mind that the P-PST is not administered until after students have been in college, and that some of the relevant content may be taught during their college program. (Yalow & Popham, 1983a, p. 11)

IOX Assessment Associates concluded that "estimates by educators in Texas suggest that students in college-preparatory programs in Texas public schools have been given substantial instruction to prepare them for all sections of the PPST" (Popham & Yalow, 1983, p. 9).

Serious questions about these conclusions need to be considered. The items, to which panelists responded, specifically asked for an opinion on "opportunity to learn." The conclusion suggests that students were "given substantial instruction." However, there is a difference between what may be available to learn and actual instruction provided. This also makes the content appropriateness of the PPST for measuring entry level skills questionable. The instructions to the panelists above imply that relevant content from the PPST is taught in college programs which, in fact, is part of teacher training programs (Yalow, 1983).

The measure of "adequacy-of-preparation" based on expert judgments had additional problems. The panels of evaluators and appraisers were sometimes unaware of the content of the PPST, of its origin, and of its difficulty level. In evaluating the appropriateness of content, judgments on difficulty or cultural bias were overlooked. ETS personnel talked with the writer about concerns of cultural bias reported on the PPST which would impact negatively on minority populations. Additionally, an interview with a representative of IOX Assessment Associates who presented the justification for content validity revealed that

knowledge about PPST development in terms of content and applicability to the guidelines of the Standards for Educational and Psychological Testing (1985) were not important considerations in the Texas process. In fact, the report clearly indicates that the validation process was designed to insure legal conformity.

Similarly, flaws seem to exist in the review of appropriate secondary texts. It must be remembered that the PPST was developed from questions and specifications taken from the NTE Core and ETS test files which were designed to test graduates at the completion of their teacher education programs. The effort to compare content in secondary school textbooks with the test specifications and PPST Skills Form 3EPS, (which was the first form of the PPST used by ETS and the first form used initially in Texas according to Popham & Yalow, 1983), seems to have limited meaning for the following reasons:

1. PPST test specifications for the NTE Core areas Reading, Writing, and Mathematics were used. These are very broad general statements such as "understands how numbers behave" and "understands and uses numbers in an appropriate way to quantify thinking," making it impossible to ascertain which actual skills would be assessed on the PPST.
2. The assumption that sufficient PPST preparation would be provided in the secondary schools is questionable. Even though some of the skills assessed on the PPST had their beginning in pre-college education, the PPST was not designed to cover secondary school curricula. Popham and Yalow (1983) indicate the PPST as a derivative of the NTE Core, which tests the content of university and teacher preparation programs.
3. Students were not provided the actual instruction needed to pass the PPST as required by previous court decisions on employment or screening. The Texas validity study provided no evidence that the content of the PPST had been "taught" in secondary schools or colleges, only that students had an "opportunity to learn" it. In fact, since content for the PPST is taken primarily from the NTE, it would probably serve better as an exit examination from teacher education programs.
4. The purpose of the textbook analysis was to gauge the amount of PPST instruction secondary school textbooks directly provide to Texas students. It was assumed that this was important because "the bulk of the preparation provided for the test undoubtedly occurs during secondary school" (Popham, Yalow, & Appel, 1983, p. 1). However, the difficulty in conducting

an adequate textbook comparative analysis and the compromise "solutions" are described by IOX Assessment Associates as follows:

In some instances, therefore, it was impossible to discern with sufficient precision, for purposes of textbook analysis, the nature of the test items circumscribed by the P-PST test specifications. No criticism of the P-PST's test specifications is intended here. For the purposes for which ETS used them, the specifications may be quite adequate. Nonetheless, those specifications proved to be insufficiently delimiting to permit matching textbook content to the specifications.

The position adopted for the present textbook analysis study was that textbook material would be considered relevant to the P-PST if, having mastered that material, a student would be able to respond correctly to corresponding P-PST test items. (Popham, Yalow, & Appel, 1983, pp. 3, 5)

Insufficient "specificity" in the PPST specifications to allow educator panelists to "guess" with reasonable accuracy is of additional concern. Accurate interpretation is thus doubtful (Holmes, 1985; Garcia, 1971).

5. Educator panelists were asked to respond to prepared questions which had a skewed tendency to provide affirmative responses in support of the use of the PPST. The response opportunities of educator panelists were interpreted as an application of the 51% rule used (Yalow & Collins, 1985). Educators were asked the following:
  - A. Do students in the college-preparatory program in your district have the opportunity to learn the content of this item?
  - B. To select one of four statements "almost all," "most," "a small amount," or "almost none".  
(Popham & Yalow, 1983)

Careful analysis of these questions is required. This is especially important since this same format has been used repeatedly in Texas to establish validation of the PPST. The "opportunity to learn" question was interpreted to mean that students in programs in the Texas public schools "had been given substantial instruction," however, this interpretation is suspect. The categories "Almost all" and "most," are contrasted with "a small amount" and "almost none." It would have been unlikely that representatives of the educator panels would have selected "small amount" or

"almost none" given the all inclusive stated specifications of the PPST. Judgments had to be made on skills taught over part of the entire educative experience. A forced choice most likely resulted in an "almost all" or "most." This process is similar to that used with the NTE and "approved" legally in U.S. v. South Carolina (1978) Other court decisions were based on high school graduation requirements which are unrelated to the use of the PPST as a screening device applied at the beginning of the professional education training component of the junior year (60 semester hours) in Texas.

Through personal interviews, the writer gathered the following information from students who took the test; from professors and administrators in universities, colleges, schools, or departments of education; and at conferences where the PPST was discussed by experts:

1. The test seems to be strongly oriented toward higher level reading skills. Students complained about not having enough time to complete the reading section.
2. In some cases, students (e.g., those from Southern Methodist University) reported passing a section of the three-part test on one administration then scoring considerably lower or failing the same section on a subsequent administration.
3. Studying for extended periods did not necessarily help students achieve a higher score.
4. Preselecting students (choosing or advising some students not to take the test based on a trial test) resulted in a higher percentage of students passing the test because this reduced the "pool" of teacher education applicants. This process simply excluded students who have been pre-tested and will likely fail the examination.
5. Some students with excellent GPAs failed the examination.

### Content Validity

According to IOX Assessment Associates and the American Psychological Association, "content validity reflects the extent to which 'the behaviors demonstrated in testing constitute a representative sample of behaviors to be exhibited in a desired performance domain'" (Popham & Yalow, 1983, p. 4).

Popham and Yalow (1983) state:

. . .the link between P-PST content and the content of teacher education programs was examined. This was referred to as the P-PST's program relevance. . . .

[Also,] an attempt was made to identify the extent to which P-PST content matched the content called for during day-to-day public school teaching in Texas. This was referred to as the P-PST's job-relevance. (p. 19)

The writer's review of the NTE and PPST indicates that PPST content is at a level of difficulty corresponding to completion of teacher education programs. ETS literature and content verification by ETS staff substantiated this finding. Thus, the PPST is more suitable for evaluating graduates from teacher education programs, as used in Delaware, than it is for screening entry into teacher education programs. In fact, this difficulty level is inappropriate for screening applicants into programs. (In Texas, students can enroll in only six semester hours of professional education courses prior to passing the PPST). Additionally, the content validation process of the PPST has ignored the performance domain. Various levels of reading, writing, and mathematics, were assessed without reference to specific behaviors. Paper-and-pencil tests simply do not measure observable behavior. Program relevance for the PPST had already been established by ETS in its development because it was composed from items in the NTE Core and ETS test files (M. Goodison, personal communication, April 4, 1985). However, to try to establish content validity for the PPST as it relates to teacher preparation programs seems inappropriate and is essentially unrelated to its current use in Texas as a screening device for admission to teacher education.

#### Content Validity in terms of Job Relevance and Program Relevance

The questions posed to faculty and educator panelists to establish program and job relevance were as follows:

1. Does a student need to know the content of this item in order to perform successfully as a teacher in Texas?
2. Does a student need to know how to write an acceptable essay in response to this assignment in order to perform successfully as a teacher in Texas? (Popham & Yalow, 1983)

Again, a forced-choice response format was used. Panelists were to respond to the first question with a "yes" or "no" for each test item and to the second with a single "yes" or "no." Panelists also made total test judgments as to whether "almost all," "most," "a small amount," or "almost none" of the PPST content must be known for successful student performance in the institution's teacher

education program (Popham & Yalow, 1983, p. 21).

It would have been quite unusual for any significant number of respondents to indicate a "small amount" or "almost none". All reviewed items came from the NTE question files at ETS and were already identified as teacher education related. The process used to establish content validity used by IOX Assessment Associates called for 51% "yes" agreement on any items or test section.

This process is not supported by "job performance" research and is not verifiable. The methods used to infer job relevance are not supported by the review of literature or by educational practice. The inferences drawn from this process constitute educated guesses at best. Decisions based on these inferences about successful performance in the programs are highly speculative for the following reasons:

1. A review of PPST tests by the writer and three other professionals with a total of over 50 years of teaching experience agreed that the questions in all three parts of the test had little relationship to the skills required by teachers in K-4 teaching assignments. ETS representatives have also indicated that the PPST is a lesser indicator of skills required at the K-4 level.
2. The writer found no research data which supports on-the-job performance assessment using the PPST or similar test results.
3. The PPST as an indicator of skills needed by teachers is probably more applicable as an exit examination from a teacher education program in accordance with the NTE and ETS guidelines. An examination of NTE and PPST test specifications indicates that they are essentially identical.
4. Use of expert judgments to predict job performance is not supported by empirical research. Experts' judgment of "job relevance" fails to comply with legal job performance evaluation criteria.
5. Yalow and Popham of IOX Assessment Associates (1983a) indicate:

Content validity was reflected, first, by the test's job relevance, that is, its relationship to the information that must be known by Texas teachers. Content validity was also indicated by the test's program relevance, that is, its relationship to the content of Texas teacher preparation programs. (p. 28)

The content validity of the PPST in Texas should be

based primarily on what is taught in the Texas teacher preparation programs. However, the PPST adoption proposes to test students on course material and content they have yet to study and learn in these same programs designed to assess exiting teacher education graduates. The PPST is probably inappropriate and unfair when used as an entrance examination into teacher education programs.

6. Even the Texas appraisal studies conducted by Yalow and Popham (1983a) suggest the need for caution:

It is difficult to render definitive conclusions regarding the suitability of the P-PST or the reasonableness of the performance standard estimates supplied for its use in Texas. Unfortunately, the PPST has not been studied in this manner in other settings, so no comparative data are available. The closest data which bear on the sorts of issues considered in the present study are those drawn from studies of the National Teacher Examinations conducted by ETS and others (pp. 32-33)

Popham and Yalow (1983) indicate additional concerns regarding implicit comparisons of PPST and NTE use:

It should be noted . . . that there are important differences between the settings in which the NTE was studied and the Texas setting in which the P-PST was appraised. In the NTE settings, the National Teacher Examinations are generally used as exit tests, that is, examinations used at the close of a preparation sequence to certify requisite professional competence. The P-PST, on the other hand, is to be used in Texas as a screening examination, that is, an examination which must be passed before a student is even admitted to a teacher education sequence. One effect of this difference is that instructional responsibility for the NTE resides rather directly in teacher preparation programs. For the P-PST, however, a far more diffuse instructional responsibility lies, one supposes, in Texas public schools. But, unlike teacher education faculty who generally recognize the existence and importance of the NTE, Texas public school educators are, for the most part, unfamiliar with the P-PST and its significance. (p. 17)

7. Although the PPST appraisal process could have withstood legal scrutiny in 1983, there is increasing

concern that test validation procedures adhere to the procedures set forth in the 1984 settlement of Golden Rule v. Washburn and to the Standards for Educational and Psychological Testing adopted in 1985 by the American Psychological Association.

8. The adoption of PPST use as the sole criterion for admission into teacher education programs may not be supported by previous court decisions. The South Carolina case was based on the NTE Common Exam. The PPST, instead, was based on the new NTE Core adopted in 1982. The Florida high school graduation test cases required that adequate preparation be supplied prior to testing. In Texas, it is questionable that preparation has, in fact, been provided to those students required to take the test.
9. Yalow and Popham (1983c) warn that raising scores on the PPST should be done with awareness of the prospective supply of candidates. The impact, especially for minority candidates, must be carefully monitored.

#### Performance Standards

IOX Assessment Associates' procedure for setting performance standards was a decidedly judgmental enterprise.

In general, two types of data have proved most helpful to standard-setters. The first of these is preference data, for example, expert judgments regarding the level at which qualified examinees are expected to perform. The second involves performance data, that is, actual test results of examinees who are similar to the individuals for whom standards must be set. . . . Results . . . will supply Texas policymakers with both the preference data and the performance data necessary to reach an enlightened decision regarding defensible passing standards for the PPST. (Yalow & Popham, 1983a, pp. 5-6.)

The process of making inferences about actual performance from "expert judgments" in order to set cut-off scores is a questionable practice (Garcia, 1971) and does not conform to the Standards for Educational and Psychological Testing (1985) or ETS guidelines. Note that the "performance data" described by IOX Assessment Associates refers to test results of examinees who are similar to individuals for whom standards must be set. It is important to note that it is not actual on-the-job performance which is a separate domain, generally unrelated to paper-and-pencil tests (Ebel, 1977). Although paper-and-pencil tests may be legally defensible in terms of earlier court cases, legal challenges will bring existing

tests into close scrutiny.

The question asked faculty panelists was "Should a student be required to know the answer to the item in order to be admitted to a teacher preparation program in Texas?" (Yalow and Popham, 1983a, p. 15). The "generic" phrase "a teacher preparation program in Texas" is crucial because of its limited meaning. Teacher education programs in Texas--as well as across the nation--vary at each institution. Only minimum requirements are mandated by state education agencies. Each institution has a designated mission statement calling for programs at that particular IHE to meet the needs of that particular student constituency. As Popham had argued earlier, faculty panelists were at a disadvantage because "the determination of passing standards is, ultimately, a judgmental enterprise. There are no standard-setting formulae available which eliminate the necessity to employ judgments in deciding what level of proficiency will be regarded as acceptable" (cited in Yalow & Popham, 1983a, p. 31).

Yalow and Popham (1983c) further describe the standard setting procedures as follows:

To provide guidance to the Board's members as they face the difficult task of setting defensible passing standards for the P-PST, a number of key individuals within Texas are being asked to supply advice to the Board regarding appropriate P-PST passing standards. Four groups of influential Texas citizens are being solicited in this effort, namely, (1) members of the Public School Professional Personnel Advisory Committee, (2) members of the Public School Board of Trustees Advisory Committee, (3) presidents of all Texas colleges and universities with approved teacher education programs, and (4) the chief administrators (e.g., deans or department chairs) of those institutions' teacher education programs. (pp. 10-11)

IOX Assessment Associates surveyed 193 individuals in Texas. The response rate was as follows: 19 from 34 members of the Public School Professional Advisory Committee, 16 from 31 responses were from members of the Public School Board of Trustees Advisory Committee, 28 from 64 were from College and university presidents, and 41 from 64 were from chief administrators of Texas teacher education programs.

The relatively low response rate of 54% (104 of 193 questionnaires returned) may be significant for a number of reasons:

1. It is probable that many respondents, like this writer at that time, were unaware of the real impact the test would have on teacher training programs and teacher supply. Additionally, many were lacking information about the PPST's appropriateness as a screening device.
2. Without appropriate knowledge of pretest score use, in their judgments, it could be possible that some of those surveyed were uncomfortable about guessing and, thus, chose not to respond.
3. It is possible that some respondents were uncertain about or unwilling to recommend cut-off scores. This especially may have been the case for educators who were aware of cut-off scores' potential impact on minority candidates. Yalow (1983) reported, "It is clear that the proportion of Hispanic and Black examinees who would pass the P-PST sections at these performance standards is exceedingly low" (p. 11).

Texas, like all other states, has the ultimate responsibility to determine how it will use the PPST. However, proper and appropriate use should be followed according to the guidelines established by ETS, the test's developer. The writer believes that the appraisal conducted by IOX Assessment Associates, the "performance" data gathered, and the proposed cut-off scores were all in contrast to the guidelines suggested by ETS. These procedures must be of ethical concern to researchers and professional educators alike.

#### Determining PPST Impact

The Mexican American Legal Defense and Educational Fund (1985) has filed a major challenge against the use of the PPST in Texas:

The challenge is in the form of a motion to enforce the court order established under U.S. v. Texas. Plaintiffs claim that Texas violates the court order, denies minority college and public school students equal protection, denies all prospective teacher education students due process and violates Texas' agreement with the Office of Civil Rights not to discriminate in higher education programs at public colleges and universities. (p. 1)

It is anticipated that PPST testing in Texas will have a serious impact on the numbers of Hispanic and black teachers seeking and gaining admission into teacher education programs. Major disparities between racial groups already exist. Table 6 reports ethnicity of Texas educational personnel for 1983-84 (see Appendix L for complete data).

Table 7 reports the ethnicity of public school students in Texas from 1981-83.

Table 6

Ethnicity of Texas Full-Time Educational Personnel for 1983-84

(Adapted from data prepared for the Texas Education Agency.)

Personnel Category	<u>Percentage of Personnel</u>		
	White	Hispanic	Black
Teachers	77	11	10
Administrators	78	12	8
Aides and Secretaries	61	27	9
<u>Total Personnel</u>	<u>74</u>	<u>14</u>	<u>10</u>

Table 7

Ethnicity of Texas Public School Children from 1981-83

(Adapted from Popham and Yalow, 1983.)

School Year	<u>Percentage of Children</u>			
	White	Hispanic	Black	Other
1981-82	56	28	15	1
1982-83	52	29	16	3

Comparison between the tables indicates existing discrepancies documenting the under-representation of Hispanic educators in Texas. It is feared that the PPST will intensify this problem by reducing the numbers of ethnic teachers available (see pass/fail rates of PPST test-takers by ethnicity from data provided by the Texas Education Agency in Appendixes M, N, O, and P).

The Texas Education Agency collects data on those teachers entering the profession who are required to take the PPST. In March 1985, 37.6% of those tested were juniors, 16.9% were seniors, 13.5% were graduates, and 1.2% failed to indicate (see Appendix Q for complete data). All graduate teachers seeking employment in Texas are required to take the PPST for employment. This "loading" with an "exited" population will create a false percentage increase in passing scores as population groups change from the sophomores and beginning juniors of the original group. Higher percentage pass rates in March, 1985 can be attributed primarily to the following:

1. Advanced students, including graduates, are now required to take the PPST.
2. Students are preselected by pretesting.
3. A high percentage of students are not electing to retake the PPST after failing it.

#### PPST Results to Date

According to the Texas Education Agency, the PPST has been administered in Texas five times since March, 1984. Table 8 provides data for each time the test was attempted by those 2,722 students who first took the PPST in March of 1984. It should be noted that the Table 8 is cumulative and, for each testing date, deals only with those students who also participated in the March, 1984 administration. Projections of pass/fail rates should not be made on any other than the first administration of the PPST because the test-taking population has changed. Goodison (1985) of ETS warns, ". . . generalizations to any specified groups other than the PPST first-time test-takers should be avoided" (p. 11).

Table 8

Cumulative Passing Rate for Students Who First Took the PPST in March 1984  
 (Source: Texas Education Agency, Austin, Texas.)

N - 2722

STATUS	FIRST ATTEMPT	SECOND ATTEMPT CUMULATIVE %	THIRD ATTEMPT CUMULATIVE %	FOURTH ATTEMPT CUMULATIVE %	FIFTH ATTEMPT CUMULATIVE %
PASS	52.9	62.2	65.9	66.6	66.6
FAIL	47.1	18.3	6.4	1.5	0
NO RETAKE	N/A	19.6	27.7	31.9	33.4
RETAKE PASS RATE	N/A	34	37	32	
CUMULATIVE	N/A	N/A	47	50	50

77

An analysis of Table 8 indicates:

First Attempt:

- A. 52.9% of all examinees passed the PPST.
- B. 47.1% of all examinees failed the PPST.

Second Attempt:

- A. 9.3% of those who had initially failed, passed the PPST, raising this cumulative pass rate to 62.2%.
- B. 18.3% failed the test for a second time.
- C. 19.6% who had failed the test the first time elected not to retake it.

Third Attempt:

- A. An additional 3.7% of the original examinees passed the test on the third attempt bringing the cumulative pass rate to 65.9%.
- B. 6.4% failed the test for the third time.
- C. 8.1% who had failed the test twice declined to take it a third time.

Fourth Attempt:

- A. An additional .7% of the original examinees passed the test on the fourth attempt raising the cumulative pass rate to 66.6%.
- B. 1.5% failed the test for the fourth time.
- C. 4.2% who had failed the test three times elected not to take it a fourth time.

Fifth Attempt:

- A. No students elected a fifth attempt.
- B. The 1.5% of examinees who had failed elected not to participate in the next administration.

In summary, the data presented shows:

1. After the fourth administration, no students are retaking the test.
2. A higher percentage of students failing the test on any attempt decided not to retake it the next time.
3. 66.6% of all students had passed the test by the fourth attempt.
4. 33.4% of all students failed to pass the test after four retakes.

Special attention is drawn to the number of students passing the PPST from March 1984 through March 1985, as shown in Table 9 (see Appendix R for complete data).

Table 9

PPST Longitudinal Data From March 1984 Through March 1985  
 (Adapted from data prepared for the Texas Education Agency.)

Candidates Tested	Numbers Passing	Percentage Passing
Black	203	1.7
Hispanic	732	6.2
Other	193	1.6
White	10,680	90.4
<u>Total Candidates</u>	11,808	99.9

After five administrations of the PPST in Texas it is clear that the PPST is having a devastating effect on ethnic minorities. Of the total number of students passing the PPST in Texas, 90% are White, 6% are Hispanic, 2% are Black, and 2% are classified as "Other." It should be noted that comparison of pass/fail rates within each ethnic group are deceptive because actual numbers of examinees from each ethnic group must be compared to the total population in order to show who is passing the examination and will make up the teaching force of tomorrow. With large numbers of teachers coming to Texas from the northern states, the percentage of Hispanics and Blacks who will become the teachers of the future will continue to decrease.

Popham and Yalow (1983) and Smith (in press) already place the minority public school enrollment in Texas between 46% and 50%, highlighting the problem of fewer ethnic minority teacher models in the schools.

#### Testing of Current Educators with the TECAT

The PPST is not the only test administered to Texas Educators. The Texas Examination of Current Administrators and Teachers (TECAT) was mandated by House Bill 72 for all professionally certified educators in Texas. This includes all teachers, administrators, counselors, librarians, and all other professional support personnel. ". . . TSTA [Texas State Teachers Association] is preparing to file a lawsuit to block the TECAT [however,] TSTA is resigned to the fact that the tests will be given as scheduled"

(Stolarek, 1985c, p. 9A). Educational Testing Service has begun refusing the use of their examinations on a limited basis to states with intentions of using their tests for purposes inconsistent with their intended uses. According to Nolan Wood, Texas Education Agency director of teacher assessment, TEA had looked at the possibility of using other tests such as the GRE, SAT, or some other national standardized tests (Stolarek, 1985d). However, "Wood said this idea was quickly put to rest because the largest testing organization, the Educational Testing Service, said they would prohibit Texas from using any of its tests for a teacher competency exam" (p. 11A).

The TECAT will consist of nine equidifferent forms from approximately 1,000 test items. The test items will reportedly be reviewed by the TECAT advisory committee for cultural, racial, or sexual bias. The State Board of Education is expected to approve performance standards or establish cut-off scores for the TECAT early in 1986.

Individuals who fail the first test will be notified directly by the test contractor and will be automatically rescheduled for a second testing opportunity. School districts will be notified of those people passing the examinations.

For \$25.00 to \$30.00 per person, or a minimum fee of \$600.00 for a six hour workshop session, the extension division of the University of Texas at Austin will provide a review course for the TECAT ("The Dread TECAT", 1985). According to Texas Commissioner of Education, W. N. Kirby, few teachers will have to worry about re-taking the test. Dr. Kirby stated, "We have many excellent teachers in Texas and we are confident the vast majority of them will pass the basic skills exam with flying colors" (Lozano, 1985, p. 17).

In an exclusive interview with the McAllen Monitor (Stolarek, 1985b), Kirby said:

. . . the legislature said we will give a one-time test, and at least we will identify those people who cannot read and write with sufficient skills to do the job, and they will, in fact, be eliminated.

What you have to understand is that if we have more than 200,000 educators in this state, and we are talking about five percent who are not competent, then you are talking about 10,000 people and that's a tremendous number of people who should not be turned loose on children . . . Therefore, the need for the one-time test.

What the teachers must do now is recognize that this was caused by those few teachers who simply don't

have the skills and shouldn't be in there. (p. 10A)

## Conclusions

1. Testing has become part of a political movement with agencies such as American Federation of Teachers and the National Education Association in the reform movement. Changes made by state agencies have been in response to educational reform through testing.
2. Each state perceives itself as having a mandate, to some degree or another, to test teachers to insure accountability in the teaching profession. As a consequence, national reform movements have become state policies which are copied from one state to another.
3. Some teachers who have failed teacher examinations are, as a consequence, perceived by the public as being incompetent. Consequently, legal and ethical issues have surfaced across the nation in response to the massive testing movement. The content of tests used to eliminate teachers from the profession are very dissimilar from legal justification indicated by the courts in the past half century. The misuse of tests have damaged the important element of fairness in the reform movement. In order to be legal, a test must be fair and equitable.
4. The testing mandate is not supported adequately by research. Scientific and ethical problems continue to loom over the testing movement. Tests designed for specific purposes are often used to measure several domains raising continued concern about appropriate test use.
5. The test guidelines recommended by test developers have often been disregarded by users. Greater responsibility rests with test users and developers to insure that tests are used ethically and fairly.
6. Test developers, who are in the business of selling tests, will most likely validate their own tests by using their own validation procedures. Court decisions have given test developers and test validators great latitude to deviate from professionally acceptable test development standards published by the American Psychological Association, Standards for Educational and Psychological Testing (1985).
7. Tests currently used for screening practicing and prospective teachers have no predictive validity. There is a clear difference between what tests measure and job-related behaviors that are measureable. There is a

clear discrepancy between what tests appear to measure and what they are perceived to measure.

8. Paper-and-pencil tests have not been shown to be related to teacher performance. Tests are limited to what they can measure. Too much reliance has been placed on a single test score in making major decisions about teacher competence. Teacher made tests and observation have more teacher confidence than standardized tests. However, a standardized test is cheap to administer avoiding costly and time consuming practices that are more subjective.
9. The appropriate use of a test can easily be misinterpreted by the public. Legislators have moved quickly to mandate testing reforms without a clear understanding of testing.
10. The underrepresentation of ethnic minorities in the teaching profession cannot be solved with mandated testing. In close association with socio-economic status, language and cultural barriers are convincing elements associated with population groups experiencing higher failure rates on tests.
11. The future is bleak for minority teachers. There is a serious disparity between the numbers of ethnic minorities in the teaching profession and Whites. Ethnic minorities are systematically being screened out of the teacher profession. Hispanics teachers are represented in lesser percentages than Black teachers.
12. High failure rates in tests are a sure sign of unfairness in test application. More time is needed to study tests and their implications. Educators do not know enough about testing.
13. Not all states using tests collected data on minority pass/fail rates. This kind of data is needed in order to make remediation possible. The states which do report this data indicate that there is a high minority failure rate. This creates a need for remediation.
14. Entry into teacher education programs should be weighed with several criteria. A single score on a single test should not be used for screening purposes. Test evaluations should include several domains at all evaluation stages.
15. Degreed teachers are no longer assumed to be competent. IHEs must accept the responsibility for guaranteeing competent teachers in the classroom.

16. Universities are confused about what to change in accordance with the reform movement. Teacher education programs are different from state to state. College programs have not changed significantly in response to the reform movement, even with state mandates.
17. Tests like the NTE that measure knowledge learned from teacher education programs should not be used as screening devices into teacher education programs. Greater responsibility for appropriate use rests with the state agencies, legislators, test developers, test validators, and the courts in confirmation with the professional standards, Standards for Educational and Psychological Testing (1985).
18. A major problem of shared responsibility to insure proper test use is exemplified with the use of the PPST in Texas. The PPST was validated in Texas under a standard procedure developed by ETS and applied by IOX Assessment Associates to validate the test in Texas. (The writer could not find any instance where a test could not be validated for any purpose using a developers recommended procedure.) ETS developed the PPST, IOX validated the PPST using an ETS procedure, and Texas had the responsibility for proper application. In proper sequence the projected use of the test should come first with sufficient time provided the developer to create the test. The second step would call for a validation process consistent with the Standards for Educational and Psychological Testing (1985) and legal procedures. Third, new examinations should be field tested several times, taking as much time as necessary, to study its impact on the intended population that will take it. Finally, a decision should be made by each state on its use being watchful of fairness in application, equity, and of assurance that quality education programs will result.

## Recommendations

1. A single cut-off score on any examination which prevents admission, continuance in any teacher training program, or for certification should be avoided.
2. States making career decisions on prospective or practicing teachers based on a single cut-off score from any test should end the practice. The myth that a single standardized examination costing \$25 can evaluate a prospective or practicing teacher should be placed to rest.
3. The use of tests should be minimized in making career decisions despite the additional costs which will become necessary to establish more acceptable and accurate testing practices.
4. Recognize that the testing movement is political in nature. That teachers, unions, lawmakers, the public, and the teaching profession can together develop the finest educational system without sacrificing equity for assumed quality.
5. In the public interest, local and state educators should work closely with state legislators in the development and approval of teacher competency testing programs. Conversely, legislators should not mandate changes which in effect "handcuff" professional educators from productive involvement.
6. Experts on testing from Institutions of Higher Education (IHEs) should be playing a major role in the enlightenment of their fellow professionals, legislators, and the public on testing appropriateness and examination use. Examinations should be reviewed by psychometric experts before adoption by states to insure that professional standards set forth by the Standards for Educational and Psychological Testing (1985) published by the APA in test development have been followed.
7. National and interstate agreement should exist on acceptable competencies which are desirable in teachers that are consistent with Standards for Educational and Psychological Testing (1985), that serves as a guide to test developers, union leaders, lawmakers and the public in mandating change.
8. Multiple criteria should be used in evaluation for entry into teacher education programs and every stage through

certification. Decisions from cumulative evidence (different domains) over a period of time will result in the best teacher selection procedure.

9. Complete data collection on pass/fail rates including ethnic minorities should be required of all states and examiners to provide information necessary to insure that informed decisions are made by lawmakers.
10. The testing mandate should make a provision for teachers with talents such as language and cultural knowledge that ethnic minorities bring into teaching which are not measured through standardized tests.
11. That teachers be required to study the applicability of testing, test development, ethical use of tests, especially as it relates to their students and themselves, in order to insure that children and adults learn the dangers of stigmatization, misclassification, and of possibly damaging people unconsciously, perpetrating a false belief about what tests measure. States should establish policies which will include ethnic minorities as a representative part of the teaching profession to insure the availability of role models consistent with the American way of life.
12. Provide adequate time for teachers who fail examinations to meet minimum requirements by relearning or remediation.
13. As a priority, and in order to be responsible to the American public, the Secretary of Education should appoint a permanent commission with broad responsibility to review and make recommendations to states on current testing practices and in examinations in use.
14. That national funding for testing research be provided to independent researchers and IHEs as opposed to developers in order to bring about a better balance and greater integrity in research practices on testing teacher competency.
15. That the legal profession become versed and knowledgeable about the application of examinations to insure that examinations are used accurately and fairly through the courts.
16. The PPST, CBEST, and NTE should not be used for admission into teacher education programs.

## References

- AACTE board meets. (1985, March). AACTE Briefs, 6(2), p. 7.
- AACTE executive committee calls for national certification. (1985, April). AACTE Briefs, pp. 1, 14, 16.
- American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association.
- Anrig, G. R. (1985a). Educational standards, testing, and equity. Phi Delta Kappan, 66, 623-625.
- Anrig, G. R. (1985b, March). Teacher education and teacher testing: The rush to mandate. Speech presented at the annual convention of the American Association of Colleges for Teacher Education, Denver, CO.
- Archer, E. L., & Leverette, H. (1985, March). Development and administration of an employee competence test. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Arciniega, T. A., & Morey, A. I. (1985, June). Hispanics and higher education: A CSU imperative (final report of the Commission on Hispanic Underrepresentation). Long Beach, CA: The California State University, Office of the Chancellor.
- Association briefs. (1985, March). AACTE Briefs, 6(2), p. 19.
- Benderson, A. (1982). Teacher competence. Focus. Princeton, NJ: Educational Testing Service.
- Bennett: Add 3Cs to the 3Rs. (1985, April 1). Education USA, 27, pp. 238-239.
- Biemiller, L. (1985, May 8). Educational Testing Service attacks author of book that calls for abolition of SAT. The Chronicle of Higher Education, p. 3.
- Board adopts NCETE report, pledges support for programs. AACTE Briefs, 6(2), pp. 1, 4-6.
- Bowen, E. & Booth, C. (1985, April). Cracking the SAT code. Time, p. 61.

Bowman, H. L., & Petry, J. R. (1984, June 15). Executive report: Validation of the NTE Core Battery and Specialty Area tests and recommendation of performance standards in Tennessee. Submitted to the Tennessee State Department of Education, Nashville. Memphis: Memphis State University, Bureau of Educational Research and Services.

Bridges, E. M., & Groves, B. (1984). Managing the incompetent teacher (National Institute of Education Contract Nos. 400-83-0013 [ERIC/CEM] and NIE-G-83-0003 [IFG]). Eugene: University of Oregon, ERIC Clearinghouse on Educational Management.

Brown, G. H., Rosen, N. L., Hill, S. T., & Olivas, M. A. (1980). The condition of education for Hispanic Americans. Washington, DC: U.S. Government Printing Office.

Brown, J. F. (1985, February 6). Implications of basic skills testing for the ethnic group composition of professional school staffs. Memo to members of the Executive Committee of the Commission on Teacher Credentialing, by R. W. Watkins, Consultant, Examinations, Sacramento, CA.

Callender, J. C. (1985, April). How good is the current generation of validity generalization estimation methods? In M. J. Burke (Chair), Validity generalization as meta-analysis. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL.

Cal.[ifornia] panel urges steps to help Hispanics. (1985, July 24). The Chronicle of Higher Education, p. 2.

Campbell, R. F., Cunningham, L. L., Nystrand, R. O., & Usdan, M. D. (1980). The organization and control of American schools (4th ed.). Columbus, OH: Merrill.

Capie, W., Ellett, C. D., & Cronin, L. (1985, April). Assessing meritorious teacher performance: Reliability, decision-making and standards setting procedures. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Carlson, R. E. (1985, April). The impact on preparation institutions of competency tests for educators. In J. Rose (Chair), The assessment boomerang returns: Competency tests for educators. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL.

- Church, G. J. (1985, July 8). A melding of cultures. Time, pp. 36-39.
- Coker, H., Medley, D. M., & Soar, R. S. (1980). How valid are expert opinions about effective teaching? Phi Delta Kappan, 62, 131-134, 149.
- Coley, R. J. (1985, March). State policies affecting teacher education and certification: Status, variations, and changes. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Cooperman, S., & Klagholz, L. (1985). New Jersey's alternate route to certification. Phi Delta Kappan, 66, 691-695.
- Corrigan, D. (1985). Politics and teacher education reform. Journal of Teacher Education, 66(1), 8-11.
- Cronin, L. L., & Capie, W. (1985, April). The influence of scoring procedures on assessment decisions and their reliability. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Cross, L. H., Impara, J. C., Frany, R. B., & Jaeger, R. M. (1984). A comparison of three methods of establishing minimum standards in the national teacher examinations. Journal of Educational Measurement, 21(2), 113-130.
- Currence, C. (1984a, August 29). Shanker urges A.F.T. to move against incompetent teachers. Education Week, p. 1.
- Currence, C. (1984b, September 5). Teacher licensing is in disarray, study contends. Education Week, pp. 1, 10.
- Danzig, A. B. (1984, January). Testing blitz on minority teachers: Intentional or unintentional? Paper presented at the annual conference of the Southwest Educational Research Association, Dallas, TX.
- Dawson-Saunders, B., Paiva, R. E. A., & Doolen, D. R. (1985, April). Predicting MCAT examination scores. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- D'Costa, A. G. (1985, April). Documenting the job-relevance of certification and licensure examinations using job analysis. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Delaney, T. (1985, May 15). Schools urged to teach 'habits of the heart.' The Denver Post, p. 9.



Denver annual meeting. (1985, March). AACTE Briefs, 6(2), pp. 10-13.

Dolly, J. P., & Williams, K. S. (1985, April). Maximizing multiple choice test scores: Generalizability of testwiseness training. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Donlon, T. F. (1985, April). Legal issues in the scaling and equating of tests. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Doyle, W. (1985). Teaching as a profession: What we know and what we need to know about teaching (National Institute of Education, Contract No. NIE-G-83-0006). Austin, TX: University of Texas at Austin, Research and Development Center for Teacher Education.

Duke, C. R. (1985). Developing a writing assessment of candidates for admission to teacher education. Journal of Teacher Education, 36(2), 7-11.

Duran, R. P. (1983). Hispanic's education and background: Predictors of college achievement. New York: College Board Publications.

Ebel, R. L. (1977). The uses of standardized testing. Bloomington, IN: Phi Delta Kappan Educational Foundation.

Echternacht, G. (1983, May 4). The validity and passing standard for the Pre-Professional Skills Tests as a certification test for Delaware teachers. Princeton, NJ: Educational Testing Service.

Educational Testing Service. (1981). Developing a test. Princeton, NJ: Author.

Educational Testing Service. (1982). Test specifications: NTE Core Battery tests. Princeton, NJ: Author.

Educational Testing Service. (1983a). ETS standards for quality and fairness. Princeton, NJ: Author.

Educational Testing Service. (1983b, September 30). Executive summary report on a study of the three NTE Core Battery tests by the state of New York. Submitted to the New York State Education Department, Albany.

- Educational Testing Service. (1984a). California Basic Educational Skills Test - Information bulletin September 1984-July 1985. Berkeley, CA: Author.
- Educational Testing Service. (1984b). California Basic Educational Skills Test: Score interpretation information. Princeton, NJ: Author.
- Educational Testing Service. (1984c). Guidelines used to score PPST essays. Princeton, NJ: Author.
- Educational Testing Service. (1984d). NTE programs test users and uses. Princeton, NJ: Author.
- Educational Testing Service. (1984e). Pre-Professional Skills Test (PPST) score interpretation guide. Princeton, NJ: Author.
- Educational Testing Service. (1984f). State-mandated testing of prospective teachers: Test used and purpose. Berkeley, CA: Author.
- Educational Testing Service. (1985). NTE programs bulletin of information: 1985 special administration service edition. Princeton, NJ: Author.
- Ekstrom, R. B., & Goertz, M. E. (1985, March). The teacher supply pipeline: The view from four states. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Ellett, C. D., & Capie, W. (1985a, April). Assessing meritorious teacher performance: A differential validity study. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Ellett, C.D., & Capie, W. (1985b, April). The Teacher Assessment and Development System (MTP Form): Instrument structure, observation procedures and content validation. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- ERIC Clearinghouse on Tests, Measurement, & Evaluation. (1980, January). ERIC highlights: Racial and cultural bias in measurement instruments. Princeton, NJ: Educational Testing Service.
- Evangelauf, J. (1985, July 3). Academe, states urged to pursue new 'covenant'. The Chronicle of Higher Education, pp. 1, 9.

- Evans, G. (1985, April 17). Gains by blacks in education found eroding. The Chronicle of Higher Education, pp. 1, 14.
- Ferraro, T. (1985, July 4). Teachers endorse testing. The McAllen Monitor, pp. 1A, 3A.
- Fisk, W. R. (1984). Predicting PPST impact: Testing and remediation at the junior level. Teacher Education & Practice, 1(2), 55-59.
- Freeman, L. D., Hess, R., III, & Kasik, M. M. (1985, March). Testing teachers and the law. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Freer, M., & Dawson, J. (1985). DON'T evaluate your teachers. Phi Delta Kappan, 66, 720-722.
- Fremer, J. (1972). Criterion-referenced interpretations of survey achievement tests (Test Development Memorandum No. TDM-72-1). Princeton, NJ: Educational Testing Service.
- Friedman, S. G. (1985, March). Testing language, language dominance, and math scores of bilingual Hispanic students. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Furtwengler, C. B. (1985, April). Evaluation procedures in the Tennessee career ladder plan. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Gage, N. L. (1984). What do we know about teaching effectiveness? Phi Delta Kappan, 66, 87-93.
- Garcia, P. A. (1971). Differences between expectations of performance in beginning first graders and actual performance. Unpublished doctoral thesis, Oregon State University, Corvallis, OR.
- Garcia, P. A. (1984). Language and cognition. Unpublished manuscript.
- Garrett, A. W. (1985, April). Measuring the efficiency and effectiveness of classroom performance. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Gideonse, H. D. (1985). State education policy in transition: Teacher education. Phi Delta Kappan, 66, 205-208.

- Goertz, M. E., Ekstrom, R. B., & Coley, R. J. (1984). The impact of state policy on entrance into the teaching profession, final report submitted to the National Institute of Education, (NIE Grant No. G83-0073). Princeton, NJ: Educational Testing Service.
- Goertz, M. E., & Pitcher, B. (1985, January). The impact of NTE use by states on teacher selection (Research Report No. RR-85-1). Princeton, NJ: Educational Testing Service.
- Goodison, M. (1985, March). Testing the basic competencies of teacher education candidates with the Pre-Professional Skills Tests (PPST). Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Haines-Saine, R. (1985, April 24). Study says Hispanic Americans would suffer under budget plan. The McAllen Monitor, p. 2A.
- Hall, G. E. (1985). Beyond the looking glass: Policies, practices and research in teacher education. Journal of Teacher Education, 66(1), 2-6.
- Hamm, D. W., & Winter, P. C. (1985, March). The education entrance examination: Testing the basic skills of prospective teachers in South Carolina. In A. M. Gallegos (Chair), Basic skills testing for prospective teachers: Promises, pitfalls, and pratfalls. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL.
- Hand, C. A., & Prather, J. E. (1985, April). The predictive validity of scholastic aptitude test scores for minority college students. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Harnisch, D. L. (1985, April). Computer application issues certification and licensure testing. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Hill, S. T., & Olivas, M. A. (1980). The condition of education for Hispanic Americans. Washington, DC: U.S. Government Printing Office.

- Hoffman, J. V., & Defino, M. E. (1985, April). State and school district intentions and the implementation of new teacher programs. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Holmes, S. E. (1985, April). Test equating issues in certification and licensure testing. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Hoover, M. E. (1984). Teacher competency tests as educational genocide for blacks: The Florida teacher certification examination. Negro Educational Review, 35(2), 70-77.
- Howe, H., II. (1985). Let's have another SAT score decline. Phi Delta Kappan, 66, 599-602.
- Hunter, R. V., & Slaughter, C. D. (1980). ETS test sensitivity review process. Princeton, NJ: Educational Testing Service.
- Imig, D. G. (1984, August). Reforms overview. AACTE Briefs, 5(6), pp. 2, 4.
- Imig, D. G. (1985, March). Planning for the future. AACTE Briefs, 6(2), pp. 2, 15.
- 'Insubordinates' in. (1985, April 22). Education USA, 27, p. 258.
- Jacobson, R. L. (1985a, February 6). AFT's president proposes national exam for teachers akin to tests for professions. The Chronicle of Higher Education, p. 19.
- Jacobson, R. L. (1985b, March 20). National license exam for teachers, extra year of study gain backing. The Chronicle of Higher Education, p. 19.
- Jacobson, R. L. (1985c, June 5). National groups starting to jockey for position in movement to reform education of teachers. The Chronicle of Higher Education, p. 14.
- Jacobson, R. L. (1985d, June 12). University education deans seek elite corps of school teachers. The Chronicle of Higher Education, pp. 1, 16.
- Kelleher, P. (1985). Inducing incompetent teachers to resign. Phi Delta Kappan, 66, 362-364.

- Kidd, J. L. (1984). Crossing the rubicon: The impact of teacher testing. Teacher Education & Practice, 1(1), 27-33.
- Kirby, W. N. (1985, June 8). Request for conditional approval of an alternative examination to the Texas Examination of Current Administrators and Teachers TECAT. Letter to the Texas State Board of Education, Austin, TX.
- Langford, M. (1985, June 9). Board of education approves teacher testing. The McAllen Monitor, pp. 1A, 9A.
- Lewis, A. C. (1985). Test scores, test scores on the wall. Phi Delta Kappan, 66, 387-388.
- Lindahl, N., & Wholeben, B. (1985). Teacher education candidates and the performance of basic skills. Teacher Education & Practice, 2(1), 35-42.
- Lines, P. M. (1985). Testing the teacher: Are there legal pitfalls? Phi Delta Kappan, 66, 618-622.
- Lozano, J. (1985, July/August). SBOE establishes dates for teacher competency exam. ATPE News, 6(1), 16-17.
- Madaus, G. F. (1985). Test scores as administrative mechanisms in educational policy. Phi Delta Kappan, 66. 611-617.
- Mannan, G., Charleston, L., & Saghafi, B. (1985, April). A comparison of the academic performance of black and white freshman students at an urban commuter campus. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Masonis, E. J. (1985, January). NTE programs special administrations service. Memo to State, School, and College Officials; Educational Testing Service, Princeton, NJ.
- McCaleb, J. (1984). Selecting a measure of oral communication as a predictor of teaching performance. Journal of Teacher Education, 35(5), 33-38.
- McCurdy, J. (1985a, July 3). Cal.[ifornia] colleges urged to prepare for more Hispanic students. The Chronicle of Higher Education, p. 2.
- McCurdy, J. (1985b, February 6). Plan to curb remedial study adopted at California State U. The Chronicle of Higher Education, p. 21.

- McLarty, J., Furtwengler, C., & Malo, G. (1985). Using multiple data sources in teacher evaluation. Paper presented to the National Council on Measurement in Education, Chicago, IL.
- Medley, D. M. (1985, April). Issues and problems in the validation of teaching and teacher professional behaviors. An invited address to Division D at the annual meeting of the American Educational Research Association, Chicago, IL.
- Messick, S., & Anderson, S. (1970). Educational testing, individual development, and social responsibility. Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. ED 047 003)
- Mexican American Legal Defense and Educational Fund. (1985). College students, American GI forum, LULAC, NAACP, MALDEF and public school students file suit against Texas teacher education entry test. San Francisco, CA.
- Minority education trends: A mixed bag. (1985, July). Hispanic Business, pp. 45-46.
- National Education Association. (1982). Excellence in our schools, teacher education: An action plan. Washington, DC.
- NTE Policy Council. (1983). Guidelines for proper use of NTE tests. Princeton, NJ: Educational Testing Service.
- Oakes, J. (1985, April). Equality as commodity: Educational reform and the politics of scarcity. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Oklahoma State Department of Education. (1984). Oklahoma Teacher Certification Testing Program. Amherst, MA: National Evaluation Systems, Inc.
- Olstad, R. G., Bean, J. L., Foster, C. D., & Marrett, A. V. (1985). A validity study of NTE general knowledge component as a predictive instrument for successful student teaching (Teacher Education Research Center Report Report No. 85-1). Seattle: University of Washington, College of Education.
- Owen, D. (1985). None of the above: Behind the myth of scholastic aptitude. Boston, MA: Houghton Mifflin Co.

- Payan, R. M., Peterson, R. E., & Castille, N. A. (1984). Access to college for Mexican Americans in the Southwest: Replication after 10 years (Report No. 84-3). New York: College Entrance Examination Board.
- Pecheone, R. (1984, February). State board of education report: Explanation of the Connecticut Competency Examination for Prospective Teachers (CONNCEPT) validation. Report submitted to the Connecticut State Department of Education, Hartford.
- Peterson, R. E. (1984, April). CBEST, NTE, and other mensurations: Notes on testing would-be teachers in California and elsewhere. Talk prepared for the Spring Conference of the California Council on the Education of Teachers, San Diego, CA.
- Phillips, G. W. (1985, April). Statistical equating of direct writing assessment. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Pigge, F. L., & Reed, P. L. (1985, March). A follow-up study of teachers' needs, proficiencies, and sources of proficiencies. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Pipho, C. (1985). Testing the teaching profession. Phi Delta Kappan, 66, 597-598.
- Popham, W. J. (1984). Teacher competency testing: The devil's dilemma. Teacher Education & Practice, 1(1), 5-9.
- Popham, W. J. (1985, April). The evaluation of teachers: A mission ahead of its measures. In R. L. Linn (Chair), Issues and problems in the validation of teaching and teacher professional behaviors. An invited address at the annual meeting of the American Educational Research Association, Chicago, IL.
- Popham, W. J., & Yalow, E. S. (1983). Appraising the Pre-Professional Skills Test for the state of Texas - Final report: A synthesis of study results. Prepared for the Texas Education Agency by IOX Assessment Associates, Culver City, CA.
- Popham, W. J., Yalow, E. S., & Appel, E. M. (1983). Appraising the Pre-Professional Skills Test for the state of Texas - Report number three: Textbook analysis. Prepared for the Texas Education Agency by IOX Assessment Associates Culver City, CA.

- Pottinger, P. S. (1975). Comments and guidelines for research in competency identification, definition and measurement (Contract No. 400-75-0036). Washington, DC: National Institute of Education. (ERIC Document Reproduction Service No. ED 134 541)
- Protesting the tests. (1985, January 7). Education USA, 27, p. 145.
- Rabianski-Carriuolo, N. (1985, February 20). Teaching basic skills to college students. The Chronicle of Higher Education, p. 88.
- Raspberry, W. (1985, March 8). Flunking basic tests. The Fresno Bee, p. B7.
- Reston, J. (1985, August 3). Hispanics in U.S. gaining ground. The Chronicle of Higher Education, p. 8.
- Reutzel, D. R., & Swindle, W. C. (1984). Preservice teacher reading abilities and the potential impact of basic skills competency tests. Teacher Education & Practice, 1, 21-26.
- Richburg, K. B. (1985, February 25). Federal exams still unavailable: Affirmative action effort is stalled. The Washington Post, p. A9.
- Richburg, K. B. (1985, June 28). 10% of Arkansas teachers fail state's competency test. The Washington Post, p. A2.
- Rosenbach, J. H. (1979, April). Instrumentation for assessing culturally different: Some conceptual issues. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco. (ERIC Document Reproduction Service No. ED 177 210)
- Rosner, F., & Grandy, J. (1982). The national teacher examinations (NTE) commons revision survey: Final report of results of a national survey to review proposed specifications for the revised commons examinations of the NTE. Princeton, NJ: Educational Testing Service.
- Salganik, L. H. (1985). Why testing reforms are so popular and how they are changing education. Phi Delta Kappan, 66, 607-610.
- Sandefur, J. T. (1985, March). State assessment trends. AACTE Briefs, 6(2), pp. 21-23.

- Sanders, T., Benton, R., Kaagan, S., Simons, L., & Teague, W. (1984, January). Staffing the Nation's Schools: A National Emergency. (Available from Council of Chief State School Officers, 400 North Capitol Street, Suite 379, Washington, DC 20001).
- SAT scores of black students improve. (1985). Phi Delta Kappan, 66, 516.
- SBOE recommends development of basic literacy exam. (1985, March/April). ATPE News, 5(5), 22-23.
- Schaffner, P. E. (1985). Competitive admission practices when the SAT is optional. The Journal of Higher Education, 56(1), 55-72.
- Schlechty, P. C. (1985). A framework for evaluating induction into teaching. Journal of Teacher Education, 36(1), 37-41.
- Short, E. C. (1985). The concept of competence: Its use and misuse in education. Journal of Teacher Education, 36(2), 2-6.
- Sikula, J. P., & Roth, R. A. (1984). Teacher preparation and certification: The call for reform. Bloomington, IN: Phi Delta Kappa Educational Foundation.
- Sirotnik, K. A. (1985, March). Responsibility vs. accountability: Towards a professional teaching profession. in E. Silva (Chair), The pedagogy of profit: A critical appraisal of current proposals for educational reform. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL.
- Smith, G. P. (in press). The impact of competency tests on teacher education: Ethical and legal issues in selecting and certifying teachers. In M. Haberman (Ed.), Research in Teacher Education. Norwood, NJ: ALEX Publishing Company.
- Special panel suggests proficiency exam for prospective teachers. (1985). Phi Delta Kappan, 66, 658-659.
- Soar, R. S., Medley, D. M., & Coker, H. (1983). Teacher evaluation: A critique of currently used methods. Phi Delta Kappan, 65, 239-246.
- Stedman, C. H. (1984, March/April). Testing for competency: A Pyrrhic Victory? The Journal of Teacher Education, 35(2), 2-5.

- Steelman, L. C., & Powell, B. (1985). Appraising the implications of the SAT for educational policy. Phi Delta Kappan, 66, 603-607.
- Stein, T., & Frankiewicz, R. G. (1985, March). Functional academic skills tests validation study for inservice teachers in Houston I.S.D. In M. F. Brightman (Chair), Local school district development and implementation of teacher basic skills competency testing. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL.
- Stewig, J. W. (1985). Reaching for links that foster strength and stability. Phi Delta Kappan, 66, 640-642.
- Stoker, W. M., & Tarrab, M. (1985). The relationship between professional skills tests and American college tests. Teacher Education & Practice, 2(1), 43-45.
- Stolarek, J. (1985a, March 22). Teachers get good, bad news on testing. The McAllen Monitor, p. 16A.
- Stolarek, J. (1985b, August 4). Education commissioner discusses teacher testing, career ladder. The McAllen Monitor, p. 10A.
- Stolarek, J. (1985c, August 8). Teacher tests discussed by TEA official. The McAllen Monitor, pp. 9A, 13A.
- Stolarek, J. (1985d, August 9). Details are offered on teacher testing plans. The McAllen Monitor, pp. 11A, 15A.
- Suhor, C. (1985). Objective tests and writing samples: How do they affect instruction in composition? Phi Delta Kappan, 66, 635-639.
- Sykes, G. (1983). Contradictions, ironies, and promises unfulfilled: A contemporary account of the status of teaching. Phi Delta Kappan, 65, 87-93.
- Tanner, C. K., & Ebers, S. M. (1985, April). Evaluation of beginning teachers in a performance-based certification program. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Teacher exams: The new rage... (1985, January 7). Education USA, 27, pp. 144-145.
- Testing fears. (1985, May 20). Education USA, 27, p. 291.
- Testing wee ones. (1985, April 22). Education USA, 27, p. 258.

- Texas Education Agency. (1983). New tests for teachers, what they mean to you. Austin, TX.
- The dread TECAT. (1985, June). The Classroom Teacher, p. 7.
- Toch, T. (1984, December 5). Test organization, insurance firm settle bias suit. Education Week, pp. 1, 14.
- Up from the bottom. (1985, March 25). Education USA, 27, p. 238.
- U.S. Department of Education. (1985, January). Indicators of education status and trends. Washington, DC: Author.
- Vasquez, M. J. T. (1985). [Review of Hispanics' education and background: Predictors of college board achievement]. Journal of Higher Education, 56, 233-237.
- Wallace, P. et al. (1966). Testing of minority group applicants for employment (Research Report No. 1966-7). Washington, DC: Equal Employment Opportunity Commission. (ERIC Document Reproduction Service No. ED 128 665)
- Weis, L. (in press). "Excellence" and student class, race, and gender cultures. In P. Altbach, G. Kelly, & L. Weis (Eds.), Excellence in education: Perspectives on policy and practice. New York: Prometheus.
- Wheeler, P., & Elias, P. (1983). California Basic Educational Skills Test: Field test and validity study report. Berkeley, CA: Educational Testing Service.
- Williams, R. L. (1970). The problem of the match and mis-match in testing black children. (ERIC Document Reproduction Service No. ED 073 131)
- Williams, W. C. (1985). Effective teaching: Gauging learning while teaching. The Journal of Higher Education, 56, 320-337.
- Willie, C. V. (1985). The problem of standardized testing in a free and pluralistic society. Phi Delta Kappan, 66, 626-628.
- Wilson, A. J. (1984a). Knowledge for teachers: The National Teacher Examinations Program, 1940 to 1970. Unpublished doctoral dissertation, University of Wisconsin, Madison, WI.

- Wilson, A. J. (1984b). Questions of content validity: The evolving relationship of the National Teacher Examinations to teacher education curricula. Proposal submitted to the American Educational Research Association for the 1985 annual meeting in Chicago, IL.
- Wilson, A. J. (1985, April). Knowledge for teachers: The origin of the National Teacher Examinations Program. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Wilson, G. W. (1985, April). The Florida Master Teacher Program. In W. Loadman (Chair), A critical examination of evaluation procedures in four merit pay/career ladder plans. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL.
- Wisniewski, R. (1984). The competence muddle. Teacher Education & Practice, 1(1), 35-38.
- Woods, H. (1985, January 27). Mission superintendent pleased by Austin meet. The McAllen Monitor, p. 5A.
- Woolever, R. M. (1985). State-mandated performance evaluation of beginning teachers: Implications for teacher educators. Journal of Teacher Education, 36(2), 22-25.
- Wright, D. P. (1985, May 21). Practitioner review of CBEST specifications, content and difficulty. Memo to members of the Planning and Research Committee by R. W. Watkins, Consultant, Planning and Research, Sacramento, CA.
- Yalow, E. S. (1983). Appraising the Pre-Professional Skills Test for the state of Texas - Report number four: Student Performance Data. Prepared for the Texas Education Agency by IOX Assessment Associates, Culver City, CA.
- Yalow, E. S., & Collins, J. L. (1985, April). Meeting the challenge of content validity. In J. Rose (Chair), The assessment boomerang returns: Competency tests for education. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL.
- Yalow, E. S., & Popham, W. J. (1983a). Appraising the Pre-Professional Skills Test for the state of Texas - Report number one: Test suitability and performance standards. Prepared for the Texas Education Agency by IOX Assessment Associates, Culver City, CA.

Yalow, E. S. & Popham, W. J. (1983b). Appraising the Pre-Professional Skills Test for the state of Texas - Report number two: Student perceptions regarding adequacy-of-preparation. Prepared for the Texas Education Agency by IOX Assessment Associates, Culver City, CA.

Yalow, E. S. & Popham, W. J. (1983c). Appraising the Pre-Professional Skills Test for the state of Texas - Report number five: Standards-advisors performance standards recommendations. Prepared for the Texas Education Agency by IOX Assessment Associates, Culver City, CA.

Young, P. A. (1975). A culture-free performance test of general learning ability (Master's thesis, Naval Postgraduate School). Springfield, VA: National Technical Information Service.

Zahorik, J. A. (1984). Can teachers adopt research findings? Journal of Teacher Education, 35(1), 34-36.

Appendix A

**Table 1**  
**States Mandating Competency Assessment of Teachers — 1984**

State	Mandate	Date	Level	Skills Tested					Type of Tests	
				Admissions	Certification	Basic	Professional	Academic	On-the-job	National Std.
Alabama		X	80/81	X	X	X	X	X	X	X
Arizona	X	X	80/81	X	X	X	X			X
Arkansas	X	X	79/83		X	X	X	X	X	X
California	X		81/83	X	X	X		X		X
Colorado	X		81/83	X		X			X	
Connecticut		X	82/83	X	X	X	X	X	X	X
Delaware		X	82/85		X	X			X	
Florida	X		78/80	X	X	X		X		X
Georgia		X	75/78		X			X	X	
Hawaii**		X	84/85			X	X	X	X	
Indiana	X		84/85	X	X	X	X	X	X	
Kansas	X		84/86		X	X	X		X	
Kentucky	X	X	82/83-85	X	X	X	X	X	X	X
Louisiana	X		77/78		X	X	X	X	X	
Maine	X		84/88		X	X	X	X	?	?
Massachusetts		X	79/82		X	X	X	X		X
Mississippi	X	X	82/86	X	X	X	X	X	X	X
Missouri		X	83/84	X		X			X	
Nebraska	X		84/85	X	X	X		X	X	?
Nevada		X	84/86	X	X	X		X	X	X
New Hampshire		X	84/85		X	X			X	
New Jersey		X	84/85		X		X	X	X	
New Mexico	X	X	81/83	X	X	X	X	X	X	X
New York		X	83/84		X	X	X		X	
North Carolina		X	79/82	X	X	X	X	X	X	
Oklahoma	X		80/82		X	X	X	X	X	X
Oregon		X	82/85	X	X	X			X	
Pennsylvania		X	84/87		X	X	X	X	X	?
Rhode Island		X	80/81		X			X		X*
South Carolina	X		79/83	X	X	X	X	X	X	X
Tennessee		X	79/79	X	X	X	X	X	X	X
Texas	X		81/84-86	X	X	X	X	X	X	X
Utah		X	79/80	X		X			X	
Vermont		X	80/82		X		X	X	X	X
Virginia	X		80/85		X	X	X	X	X	X
Washington		X	78/83	X		X			X	X
West Virginia		X	82/85		X	X	X	X	X	X
Wyoming		X	82/82	X		X			X	
<b>TOTALS-38</b>	<b>17</b>	<b>26</b>		<b>21</b>	<b>32</b>	<b>34</b>	<b>25</b>	<b>26</b>	<b>13</b>	<b>28</b>

\* Vocational teachers only.

\*\* Hawaii uses competency tests for employment only, not for admission or certification.

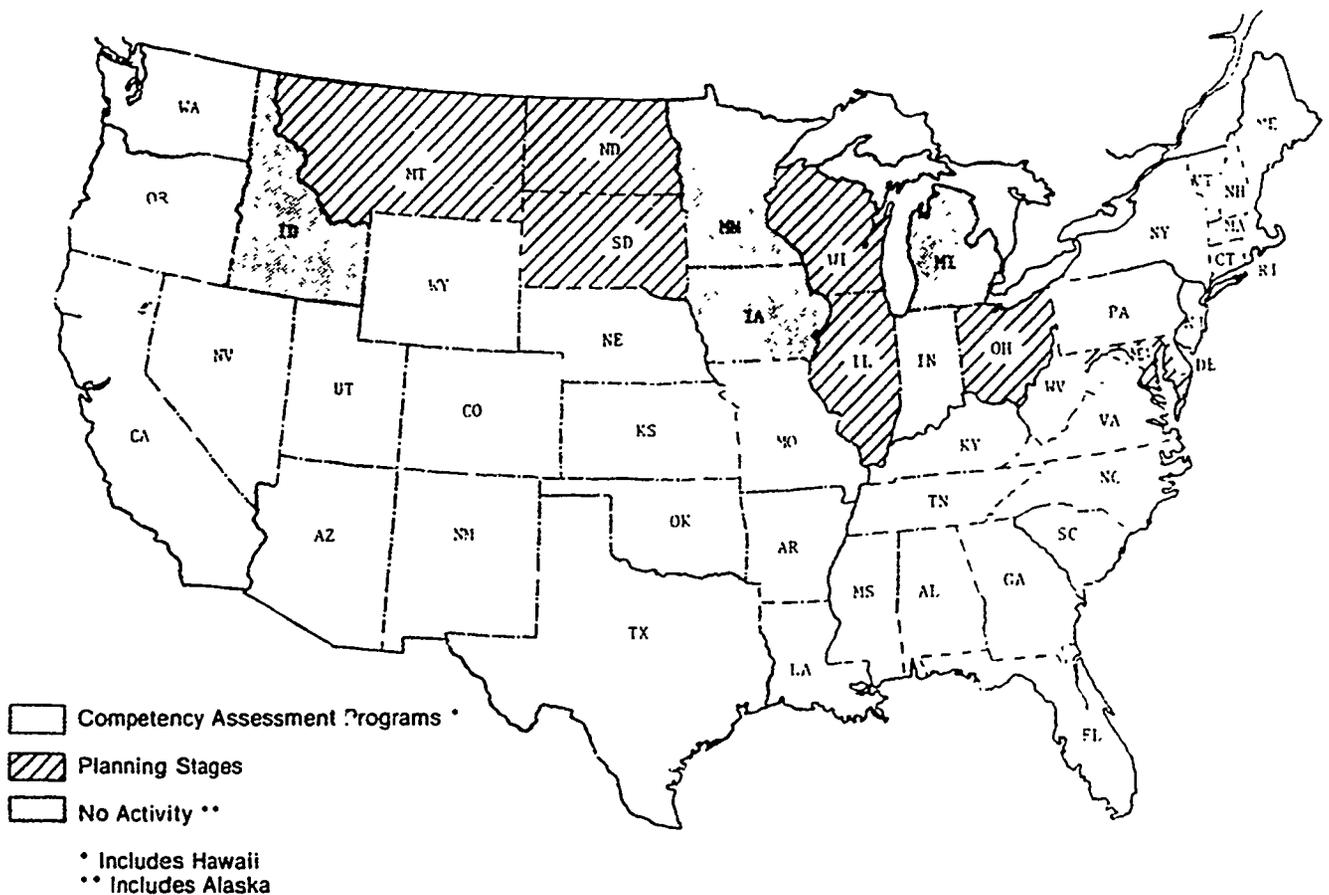
Source: Sandefur, J. T. (1985, March). State Assessment Trends. AACTE Briefs, 6(2), p. 18.

Appendix B

**Table 2**  
**States Reporting Planning or Discussion of Competency Assessment of Teachers—1984**

State	Level		Skills Tested					Type of Tests		Year First Reported Planning
	Admissions	Certification	Basic	Professional	Academic	On-the-job	Standardized	Customized		
Illinois	X	X	X							1980
Maryland		X	X	X	X	X	X			1982
Montana		X	X					X		1980
North Dakota		X	X							1983
Ohio		X	X	X	X				X	1982
South Dakota	X	X	X	X	X			X		1984
Wisconsin	X	X	X	X	X			X		1980
Totals-7	3	7	7	4	4	1	4	1		

Competency Assessment of Teachers - 1984



Source: Sandefur, J. T. (1985, March). State Assessment Trends. *AACTE Briefs*, 6(2), pp. 17, 19.

Table 3

## Appendix C

Teacher Competency Test Pass Rates  
By Ethnicity For Ten States

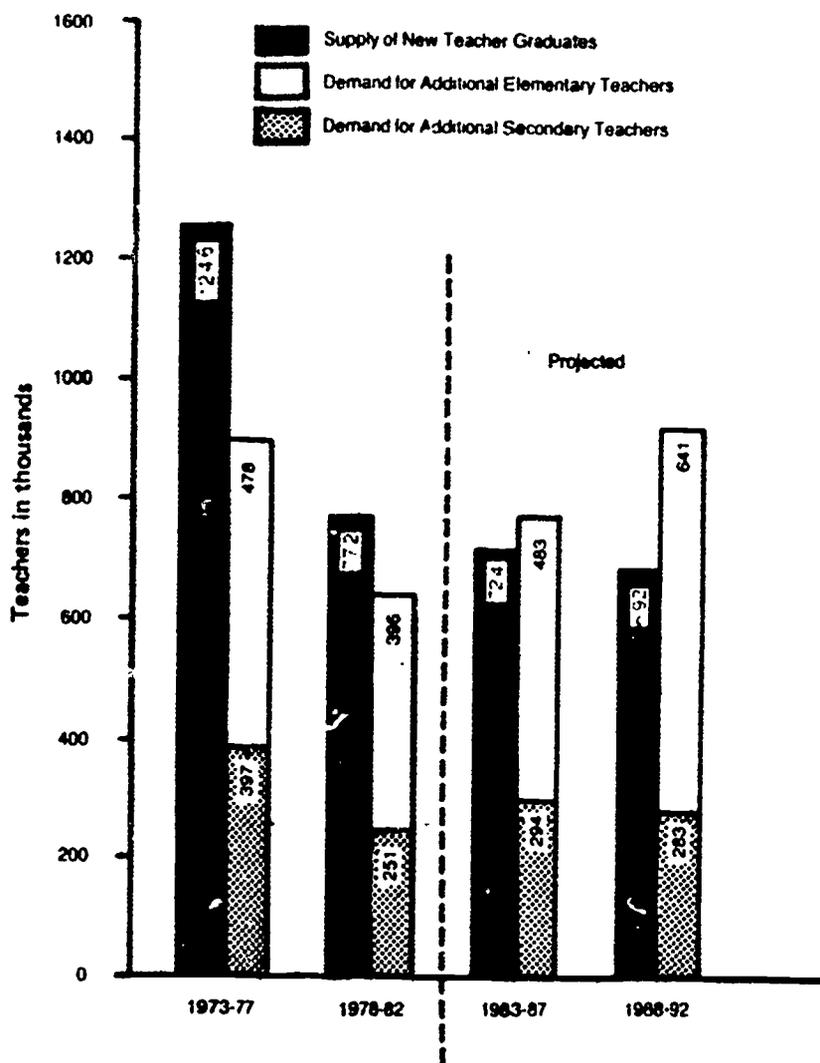
State	Pass Rates By Percent						Test
	Anglos	Asians	Blacks	Hispanics	Native Americans	All	
Alabama	86		43			81	AICT(NES)
Arizona 1/6/83	73	50	24	42	22	66	ATPE
7/9/83	70	25	41	36	19	59	
California	76	50	26	38	67	68	CBEST(ETS)
Florida 6/82	92	67	37	57	90	85	FCTE
2/83	90	63	35	51	100	84	Customized
Georgia	87		34			78	CRTCT(NES)
Louisiana	78		15			77	NTE(ETS)
Mississippi*	97-100		54-70			NA	NTE(ETS)
Oklahoma	79	82	45	71	70	78	OCT(NES) Customized
Texas	62	47**	10	19	47**	54	P-PST(ETS)
Virginia* (Trial Testing) Communication Skills	97%		56%			NA	NTE(ETS)
General Know- ledge	99%		69%			NA	
Prof. Knowledge	99%		83%			NA	

\*Pass rates at predominately white and black public institutions.

\*\*Asian and Native American candidates are reported in a combined "Others" category in the Texas reporting system.

Source: Smith, G. P. (in press). The Impact of Competency Tests on Teacher Education: Ethical and Legal Issues in Selecting and Certifying Teachers. In M. Haberman (Ed.) Research in Teacher Education. Norwood, NJ: ABLEX Publishing Company.

**Estimated Teacher Supply and Estimated Demand for Additional Teachers**



Source: NCES, *The Condition of Education*, 1984

● Beginning in the mid-1980's, the demand for additional teachers is projected to exceed the supply of new teacher graduates. Elementary schools should provide two-thirds of this demand

Source: U.S. Department of Education. (1985, January). Indicators of Education Status and Trends, p. 29.

Appendix E

11/100 12/89

TEACHERS EMPLOYED (IN FULL-TIME EQUIVALENTS) IN PUBLIC AND PRIVATE ELEMENTARY AND SECONDARY SCHOOLS, AND CANDIDATE SHORTAGES, BY FIELD OF ASSIGNMENT: 50 STATES AND D.C., NOVEMBER 1, 1983

Field of Assignment	Total Teachers		Candidate Shortages		Shortages per 1000 Teachers
	Number	Percent	Number	Percent	
Total	2,553,300	100.0	4,000	100.0	1.6
Preprimary Education	89,100	3.5	80	2.0	0.9
General Elementary Education	873,300	34.2	740	18.7	0.8
Art	50,700	2.0	180	4.6	3.6
Basic Skills/Remedial Education	42,300	1.7	120	3.1	2.9
Bilingual Education	29,900	1.2	260	6.6	8.8
Biological and Physical Sciences	131,100	5.1	230	5.7	1.7
Biology	28,800	1.1	50	1.2	1.7
Chemistry	14,600	0.6	30	0.7	1.9
Physics	8,700	0.3	40	1.0	4.5
General and All Other Sciences	79,000	3.1	110	2.8	1.4
Business (non-vocational)	53,800	2.1	20	0.5	0.4
Computer Science	9,200	0.4	30	0.9	3.7
English Language Arts	182,700	7.2	170	4.3	0.9
Foreign Languages	50,400	2.0	80	1.9	1.5
Health, Physical Education	151,500	5.1	100	2.5	0.8
Home Economics	38,100	1.5	30	0.7	0.7
Industrial Arts	43,700	1.7	80	2.1	1.9
Mathematics	147,100	5.8	260	6.6	1.8
Music	79,100	3.1	240	6.1	3.1
Reading	47,700	1.9	20	0.5	0.4
Social Studies/Social Sciences	142,400	5.6	70	1.7	0.5
Special Education	264,100	10.3	1,030	25.9	3.9
Mentally Retarded	54,400	2.1	150	3.9	2.8
Seriously Emotionally Disturbed	26,800	1.0	100	2.5	3.7
Specific Learning Disabled	73,200	2.9	190	4.8	2.6
Speech Impaired	27,700	1.1	180	4.4	6.3
Other Special Education	82,000	3.2	400	10.3	5.0
Vocational Education	64,300	2.5	70	1.7	1.1
Other Elementary Education	29,800	1.2	30	0.8	1.1
Other Secondary Education	53,500	2.1	120	3.0	2.2

Source: 1983-84 Survey of Teacher Demand and Shortage (NCES).

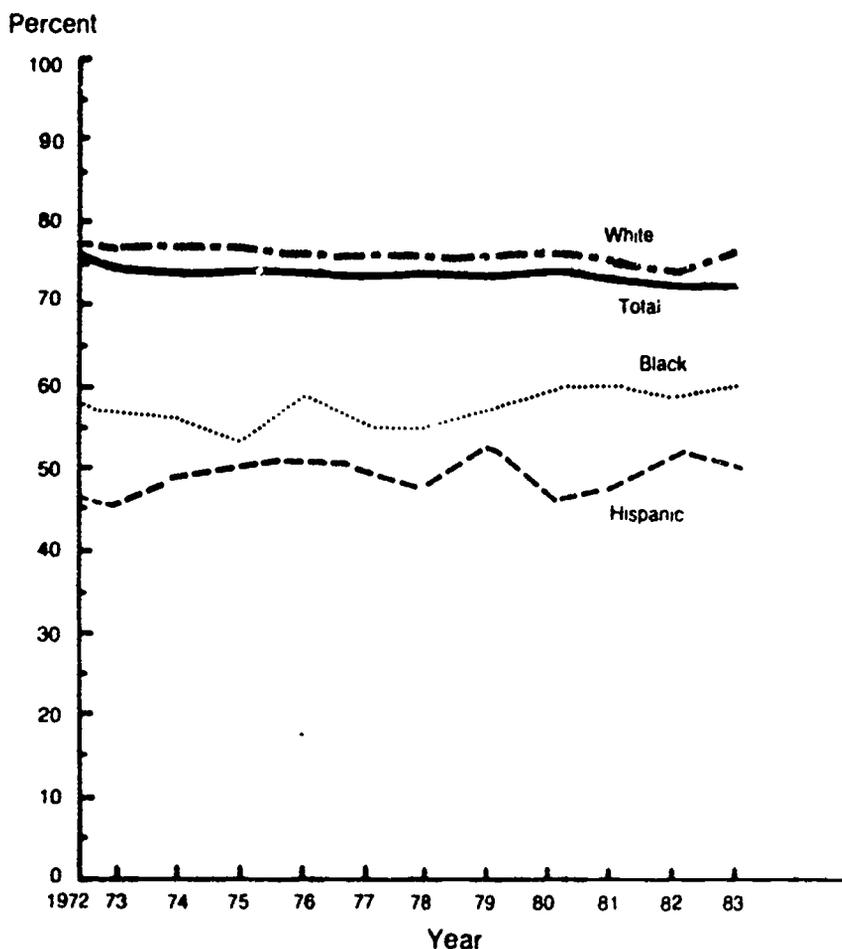
Note: Percentages are calculated on unrounded numbers.

Source: U.S. Department of Education. (1985, January). Indicators of Education Status and Trends, p. A-11.

## Appendix F

### Percent High School Graduates by Race/Ethnicity: 1972 to 1983

(Percent of 18 and 19 Year Olds Completing 12 Years of Education; totals include GEDs)

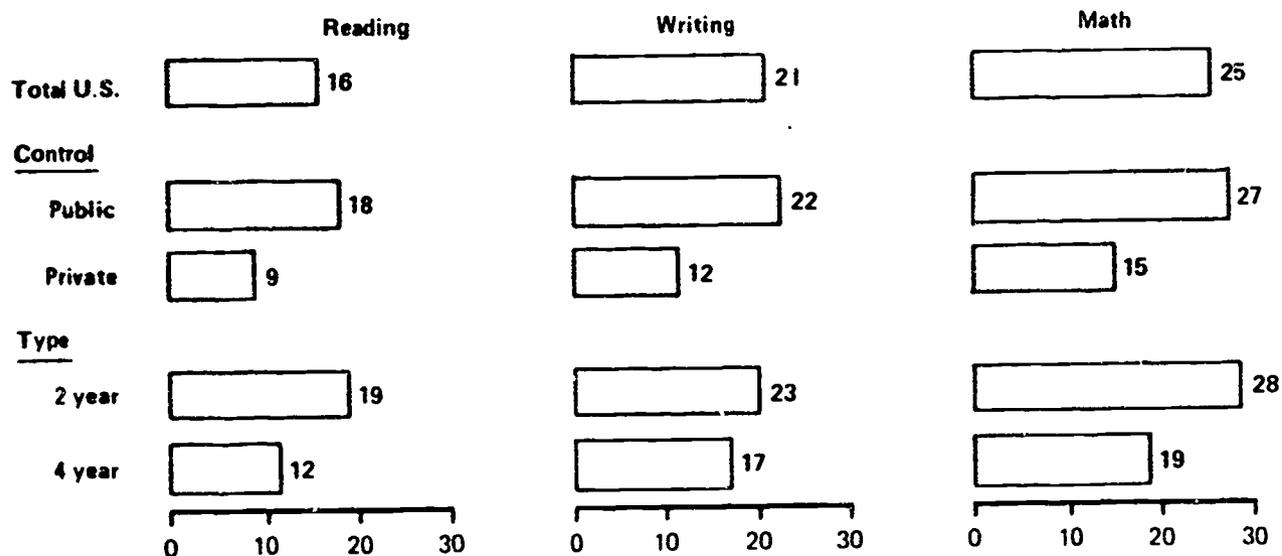


Source: Bureau of the Census Current Population Reports

- The proportion of 18- and 19-year olds who have graduated from high school declined from a peak of 74.8 percent in 1972 to 72.0 percent in 1982.
- The percentages of high school graduates for Blacks and Hispanics are several points below Whites for each year, but are slightly higher in 1982 than in 1972.

Source: U.S. Department of Education. (1985, January). Indicators of Education Status and Trends, p. 11.

Percentage of Freshmen Enrolled in Remedial Courses in Institutions of Higher Education  
by Subject: 1983-84



Source. Fast Response Survey System.

- Sixteen percent of college freshmen are enrolled in remedial reading, 21 percent are in remedial writing, and 25 percent are in remedial math. These percentages vary by control, type, geographic region, and selectivity of the college.
- 82 percent of higher education institutions with first-year programs offered remedial courses in reading, writing, or math in 1984. Of these, 63 percent had remedial enrollment increases of 10 percent or more, 33 percent had relatively stable remedial enrollment, and only 4 percent had decreases of 10 percent or more.

Source: U.S. Department of Education. (1985, January). Indicators of Education Status and Trends, p. 15.

BEST COPY AVAILABLE

Appendix H

AVERAGE SAT SCORES OF ALL ENTERING FRESHMEN  
IN INSTITUTIONS OF NORTH CAROLINA  
FALL 1980

<u>INSTITUTION</u>	<u>SAT TOTAL</u>
Duke	1253
Davidson	1215
Wake Forest	1111
UNC-Chapel Hill	1064
N.C. State	1004
Guilford	968
Salem	965
UNC-Greensboro	949
Meredith	937
Warren Wilson	937
Lenoir-Rhyne	930
UNC-Charlotte	920
UNC-Asheville	896
Appalachian	894
St. Andrews	892
* St. Augustine's	885
Queen's	878
East Carolina	874
High Point	868
UNC-Wilmington	848
Catawba	843
Pfeiffer	841
Greensboro	826
Western Carolina	821
Mars Hill	814
Campbell	809
Atlantic Christian	803
Gardner-Webb	799
Pembroke	784
Elon	780
Wingate	775
Methodist	764
N.C. Wesleyan	749
Sacred Heart	741
* N.C.A. and T.	667
* Winston-Salem State	643
* N.C. Central	624
* Livingstone	600
* Bennett	597
* Fayetteville State	587
* Shaw	584
* Johnson C. Smith	573
* Elizabeth City State	560
* Barber-Scotia	550

\* Predominantly Black institutions  
(Pattern has not changed much since 1980.)

Source: North Carolina State University, School of Education

**Trends in Total Enrollment in Institutions of Higher Education,  
by Level of Institution and Race/Ethnicity: Fall 1976 to Fall 1982**

Race/Ethnicity and citizenship	Number Enrolled				% Enrolled			
	76	78	80	82	76	78	80	82
4-year institutions .....	7,090	7,187	7,548	7,629	100.0	100.0	100.0	100.0
White, non-Hispanic .....	5,984	6,013	6,259	6,289	84.4	83.7	82.9	82.4
Total Minority .....	930	973	1,048	1,070	13.1	13.5	13.9	14.0
Black, non-Hispanic .....	603	611	633	611	8.5	8.5	8.4	8.0
Hispanic .....	173	190	216	228	2.4	2.6	2.9	3.0
Asian or Pacific Islander .....	118	137	162	193	1.7	1.9	2.1	2.5
American Native/ Alaskan Native .....	35	35	37	38	.5	.5	.5	.5
Non-resident alien .....	176	200	241	269	2.5	2.8	3.2	3.5
2-year institutions .....	3,880	4,028	4,490	4,699	100.0	100.0	100.0	100.0
White, non-Hispanic .....	3,077	3,167	3,532	3,657	79.3	78.6	78.7	77.8
Total Minority .....	761	810	894	981	19.6	20.1	19.9	20.9
Black, non-Hispanic .....	429	443	468	483	11.1	11.0	10.4	10.3
Hispanic .....	210	227	255	291	5.4	5.6	5.7	6.2
Asian or Pacific Islander .....	79	97	124	158	2.0	2.4	2.8	3.4
American Indian/ Alaskan Native .....	41	43	47	49	1.1	1.1	1.0	1.0
Non-resident alien .....	42	52	64	61	1.1	1.3	1.4	1.3
<b>Total</b> .....	10,970	11,215	12,038	12,328	100	100	100	100
<b>White</b> .....	9,061	9,180	9,791	9,946	82.6	81.9	81.3	80.7
<b>Total Minority</b> .....	1,691	1,783	1,942	2,051	15.4	15.9	16.1	16.6
Black .....	1,032	1,054	1,101	1,094	9.4	9.4	9.2	8.9
Hispanic .....	383	417	471	519	3.5	3.7	3.9	4.2
Asian-Pacific .....	197	234	286	351	1.8	2.1	2.4	2.8
Native American .....	76	78	84	87	.7	.7	.7	.7
<b>Non-resident Alien</b> .....	218	252	305	330	2.0	2.2	2.5	2.7

*NOTE. Excludes enrollment in U.S. Service Schools. Numbers in thousands; percentages may not add to 100 due to rounding.*

*SOURCE: U.S. Department of Education, Office for Civil Rights, unpublished tabulations (December 1983) and National Center for Education Statistics, Opening Fall Enrollment, Fall 1982, unpublished tabulations (December 1983).*

*Office for Civil Rights, unpublished tabulations, reported NCES, Condition of Education, 84 p. 76.*

Source: Minority Education Trends: A Mixed Bag. (1985, July). Hispanic Business, p. 46.

Appendix J  
Questionnaire

<name>

PART 1. ADMISSION INTO TEACHER EDUCATION

1. Does your state administer a test for admission into teacher education programs?

- Yes  
 No

If the answer is "No", please proceed to Part II on page 2.

2. Type of Test (check all that apply)

- National  
 State  
 Regional  
 Local  
 Custom

Explain (if necessary): \_\_\_\_\_  
\_\_\_\_\_

3. Name of Test(s): \_\_\_\_\_

4. Developed by:

- Educational Testing Service (ETS)  
 National Evaluation Systems, Inc. (NES)  
 Other (specify): \_\_\_\_\_

Explain (if necessary): \_\_\_\_\_

5. What does the examination test? (Check more than one if necessary).

- Basic Skills  
 If a student can perform in a teacher education program  
 Knowledge a student should have to do well as a teacher  
 Teacher Competence - knowledge and a set of attitudes required for satisfactory performance as a teacher  
 Teacher Performance - a pattern of behavior a teacher displays while teaching a class  
 Teacher performance based on student learning outcomes - changes in student knowledge  
 Other (specify): \_\_\_\_\_

6. The test was validated by:

- Educational Testing Service (ETS)  
 National Evaluation Systems, Inc. (NES)  
 Other (specify): \_\_\_\_\_

Survey  
Page 2

7. Is a cut-off score on the test(s) used as sole criterion for determining admission?

- Yes  
 No

Explain (if necessary): \_\_\_\_\_

8. Please indicate pass/failure rates for the various ethnic groups. Attach additional information if necessary.

% White - pass  
 % White - fail

% Black - pass  
 % Black - fail

% Hispanic - pass  
 % Hispanic - fail

% Other - pass (specify): \_\_\_\_\_  
 % Other - fail (specify): \_\_\_\_\_

9. Minorities pass the examination at:

- a lower percentage rate than Whites  
 an equal percentage rate to Whites  
 a higher percentage rate than Whites

10. Please name and describe any other examinations of this type used in your state.

PART II. PROFESSIONAL EDUCATION (PEDAGOGY)

1. Does your state administer a test for professional education (pedagogy)?

- Yes  
 No

If the answer is "No", please proceed to Part III on page 4.

2. Type of Test (check all that apply)

- National  
 State  
 Regional  
 Local  
 Custom

Explain (if necessary): \_\_\_\_\_

3. Name of Test(s): \_\_\_\_\_

4. Developed by:

- Educational Testing Service (ETS)  
 National Evaluation Systems, Inc. (NES)  
 Other (specify): \_\_\_\_\_

Explain (if necessary): \_\_\_\_\_

5. What does the examination test? (Check more than one if necessary).

- Basic Skills  
 If a student can perform in a teacher education program  
 Knowledge a student should have to do well as a teacher  
 Teacher Competence - knowledge and a set of attitudes required for satisfactory performance as a teacher  
 Teacher Performance - a pattern of behavior a teacher displays while teaching a class  
 Teacher performance based on student learning outcomes - changes in student knowledge  
 Other (specify): \_\_\_\_\_

6. The test was validated by:

- Educational Testing Service (ETS)  
 National Evaluation Systems, Inc. (NES)  
 Other (specify): \_\_\_\_\_

7. Is a cut-off score on the test(s) used as sole criterion for determining professional education?

- Yes  
 No

Explain (if necessary): \_\_\_\_\_

8. Please indicate pass/failure rates for the various ethnic groups. Attach additional information if necessary.

- % White - pass  
 % White - fail

- % Black - pass  
 % Black - fail

- % Hispanic - pass  
 % Hispanic - fail

- % Other - pass (specify): \_\_\_\_\_  
 % Other - fail (specify): \_\_\_\_\_

9. Minorities pass the examination at:  
 a lower percentage rate than Whites  
 an equal percentage rate to Whites  
 a higher percentage rate than Whites
10. Please name and describe any other examinations of this type used in your state.

PART III. ACADEMIC

1. Does your state administer a test for academic subject areas?  
 Yes  
 No

If the answer is "No", please proceed to Part IV on page 5.

2. Type of Test (check all that apply)  
 National  
 State  
 Regional  
 Local  
 Custom

Explain (if necessary): \_\_\_\_\_  
\_\_\_\_\_

3. Name of Test(s): \_\_\_\_\_

4. Developed by:  
 Educational Testing Service (ETS)  
 National Evaluation Systems, Inc. (NES)  
 Other (specify): \_\_\_\_\_

Explain (if necessary): \_\_\_\_\_

5. What does the examination test? (Check more than one if necessary).  
 Basic Skills  
 If a student can perform in a teacher education program  
 Knowledge a student should have to do well as a teacher  
 Teacher Competence - knowledge and a set of attitudes required for satisfactory performance as a teacher  
 Teacher Performance - a pattern of behavior a teacher displays while teaching a class  
 Teacher performance based on student learning outcomes - changes in student knowledge

Other (specify): \_\_\_\_\_

6. The test was validated by:

Educational Testing Service (ETS)

National Evaluation Systems, Inc. (NES)

Other (specify): \_\_\_\_\_

7. Is a cut-off score on the test(s) used as sole criterion for determining academic performance?

Yes

No

Explain (if necessary): \_\_\_\_\_

8. Please indicate pass/failure rates for the various ethnic groups. Attach additional information if necessary.

% White - pass

% White - fail

% Black - pass

% Black - fail

% Hispanic - pass

% Hispanic - fail

% Other - pass (specify): \_\_\_\_\_

% Other - fail (specify): \_\_\_\_\_

9. Minorities pass the examination at:

a lower percentage rate than Whites

an equal percentage rate to Whites

a higher percentage rate than Whites

10. Please name and describe any other examinations of this type used in your state.

#### PART IV. ON-THE-JOB

1. Does your state administer a test for on-the-job performance?

Yes

No

If the answer is "No", please proceed to Part V on page 7.

2. Type of Test (check all that apply)

- National
- State
- Regional
- Local
- Custom

Explain (if necessary): \_\_\_\_\_  
\_\_\_\_\_

3. Name of Test(s): \_\_\_\_\_

4. Developed by:

- Educational Testing Service (ETS)
- National Evaluation Systems, Inc. (NES)
- Other (specify): \_\_\_\_\_

Explain (if necessary): \_\_\_\_\_

5. What does the examination test? (Check more than one if necessary).

- Basic Skills
- If a student can perform in a teacher education program
- Knowledge a student should have to do well as a teacher
- Teacher Competence - knowledge and a set of attitudes required for satisfactory performance as a teacher
- Teacher Performance - a pattern of behavior a teacher displays while teaching a class
- Teacher performance based on student learning outcomes - changes in student knowledge
- Other (specify): \_\_\_\_\_

6. The test was validated by:

- Educational Testing Service (ETS)
- National Evaluation Systems, Inc. (NES)
- Other (specify): \_\_\_\_\_

7. Is a cut-off score on the test(s) used as sole criterion for determining on-the-job performance?

- Yes
- No

Explain (if necessary): \_\_\_\_\_

8. Please indicate pass/failure rates for the various ethnic groups. Attach additional information if necessary.

\_\_\_ % White - pass  
\_\_\_ % White - fail

\_\_\_ % Black - pass  
\_\_\_ % Black - fail

\_\_\_ % Hispanic - pass  
\_\_\_ % Hispanic - fail

\_\_\_ % Other - pass (specify): \_\_\_\_\_  
\_\_\_ % Other - fail (specify): \_\_\_\_\_

9. Minorities pass the examination at:  
\_\_\_ a lower percentage rate than Whites  
\_\_\_ an equal percentage rate to Whites  
\_\_\_ a higher percentage rate than Whites

10. Please name and describe any other examinations of this type used in your state.

#### PART V. CERTIFICATION

1. Does your state administer a test for teacher certification?

\_\_\_ Yes  
\_\_\_ No

If the answer is "No", please proceed to General Information on page 9.

2. Type of Test (check all that apply)

\_\_\_ National  
\_\_\_ State  
\_\_\_ Regional  
\_\_\_ Local  
\_\_\_ Custom

Explain (if necessary): \_\_\_\_\_  
\_\_\_\_\_

3. Name of Test(s): \_\_\_\_\_

4. Developed by:

\_\_\_ Educational Testing Service (ETS)  
\_\_\_ National Evaluation Systems, Inc. (NES)  
\_\_\_ Other (specify): \_\_\_\_\_

Explain (if necessary): \_\_\_\_\_

5. What does the examination test? (Check more than one if necessary).

- Basic Skills  
 If a student can perform in a teacher education program  
 Knowledge a student should have to do well as a teacher  
 Teacher Competence - knowledge and a set of attitudes required for satisfactory performance as a teacher  
 Teacher Performance - a pattern of behavior a teacher displays while teaching a class  
 Teacher performance based on student learning outcomes - changes in student knowledge  
 Other (specify): \_\_\_\_\_

6. The test was validated by:

- Educational Testing Service (ETS)  
 National Evaluation Systems, Inc. (NES)  
 Other (specify): \_\_\_\_\_

7. Is a cut-off score on the test(s) used as sole criterion for determining certification?

- Yes  
 No

Explain (if necessary): \_\_\_\_\_

8. Please indicate pass/failure rates for the various ethnic groups. Attach additional information if necessary.

- % White - pass  
 % White - fail

- % Black - pass  
 % Black - fail

- % Hispanic - pass  
 % Hispanic - fail

- % Other - pass (specify): \_\_\_\_\_  
 % Other - fail (specify): \_\_\_\_\_

9. Minorities pass the examination at:

- a lower percentage rate than Whites  
 an equal percentage rate to Whites  
 a higher percentage rate than Whites

10. Please name and describe any other examinations of this type used in your state.

GENERAL INFORMATION

1. Would you please return a copy of your validation report to Dr. Peter Garcia for study?

Yes

No

Explain (if necessary): \_\_\_\_\_

2. Please indicate whether you would like a copy of the final report sent to you.

Yes

No

3. Please submit names and addresses of persons you believe would be interested in receiving a copy of the final report.

4. You are welcome to include any comments that you feel would be helpful in this study. Thank you for your cooperation in completing this survey.

## Appendix K

### PASS RATES FOR 10 STATES' TEACHER COMPETENCY TESTS

	All	Whites	Asians	Blacks	Hispanics	Native Americans	
Alabama	81%	86%	---	43%	---	---	
Arizona	1/6/83 7/9/83	66 59	73 70	50% 25	24 41	42% 36	22% 19
California		68	76	50	26	38	67
Florida	6/82 2/83	85 84	92 90	67 63	37 35	57 51	90 100
Georgia		78	87	---	34	---	---
Louisiana		77	78	---	15	---	---
Mississippi		---	97 <sup>1</sup> 100 <sup>2</sup>	---	54 70	---	---
Oklahoma		78	79	32	45	71	70
Texas		54	62	47	10	19	47
Virginia (Trial Testing)							
Communication Skills		---	97	---	56	---	---
General Knowledge		---	99	---	69	---	---
Professional Knowledge		---	99	---	83	---	---

<sup>1</sup> Pass rates at predominately white public institutions.

<sup>2</sup> Pass rates at predominately black public institutions.

Source: American Association of Colleges for Teacher Education.

Source: Brown, J. F. (1985, February 6). Implications of Basic Skills Testing for the Ethnic Group Composition of Professional School Staffs. Memo to members of the Executive Committee of the Commission on Teacher Credentialing, by R. W. Watkins, Consultant, Examinations, Sacramento, CA.

Appendix L

09/20/84 14:01

T A B L E I I - 2  
TEXAS 1983-84 PPR FULL-TIME EQUIVALENT PERSONNEL (FTE)  
TOTALS FOR EACH GENERAL PERSONNEL TYPE, BY SEX AND ETHNICITY

Personnel Category	M A L E						F E M A L E						Grand Total
	White	Hispanic	Black	Asian	Amer Indian	Male Total	White	Hispanic	Black	Asian	Amer Indian	Female Total	
TOTAL PERSONNEL	40,994	7,879	5,006	114	70	54,063	139,574	27,813	20,198	398	237	188,221	242,287
TEACHERS	30,917	5,610	3,815	86	53	40,480	100,321	14,524	14,199	283	134	129,461	169,943
NO DEGREE	634	141	47	1	1	825	183	81	11	1	0	276	1,100
BA DEGREE	16,593	3,835	2,023	36	29	21,516	66,911	10,650	7,163	171	82	84,987	107,505
MA DEGREE	13,422	1,618	1,716	41	23	16,820	33,007	3,769	6,988	106	52	43,922	60,742
ADV DEGREE	268	15	29	8	0	320	219	14	38	4	0	276	595
SUPPORT & ADMINISTRATIVE	9,763	1,571	978	8	17	12,336	11,949	1,575	1,744	30	14	15,313	27,649
SUPPORT	1,774	406	204	3	2	2,389	9,354	1,108	1,352	19	9	11,842	14,231
SUPERVISORS	299	72	27	1	1	400	848	163	138	1	2	1,158	1,500
COUNSELORS	1,263	288	140	2	1	1,694	3,830	507	614	5	4	4,959	6,653
LIBRARIANS	130	10	6	0	0	146	2,756	137	326	7	1	3,227	3,373
HEALTH SERVICES	11	3	1	0	0	14	1,717	229	222	4	2	2,174	2,188
THERAPISTS	9	5	1	0	0	15	123	27	9	2	0	161	176
OTHER	62	28	30	0	0	120	80	40	42	0	0	162	282
ADMINISTRATIVE	7,989	1,165	774	4	15	9,948	2,595	467	392	11	6	3,471	13,418
ADMIN OFFICERS	2,366	388	184	2	3	2,943	1,113	189	129	10	2	1,443	4,386
FULL TIME PRINCPL	3,326	443	327	1	8	4,105	841	166	163	1	3	1,175	5,280
PART-TIME PRINCPL	1,247	280	257	0	2	1,785	513	98	94	0	0	705	2,490
SUPERINTENDENTS	986	41	3	0	2	1,032	16	0	0	0	0	16	1,048
OFFICERS: NO DEGR	65	14	3	1	0	82	112	13	6	0	0	132	214
AIDES & SECRETARIES	314	699	213	21	0	1,247	27,304	11,714	4,254	86	89	43,448	44,695
AIDES	186	512	163	20	0	880	12,371	7,897	2,781	59	55	23,163	24,044
SECRETARIES	128	187	50	1	0	367	14,933	3,817	1,473	27	34	20,285	20,652

Source: Texas Education Agency, Austin, Texas.

123

PERFORMANCE BY NUMBER OF TESTS PASSED  
March 1984

State Totals	Number Tested	Passed Three Tests	Passed Two Tests	Passed One Test	Passed None	Note
TAKING OUT PARTIAL RECORDS (51 took only 1 or 2 tests)	2,687	1,440 or 54%	503 or 19%	365 or 13%	379 or 14%	
HISPANIC	436	84 or 19%	84 or 19%	96 or 22%	172 or 39%	16 took only 1 or 2 tests. 2 of these passed all taken.
BLACK	126	12 or 10%	19 or 15%	23 or 18%	72 or 57%	7 took only 1 or two tests. 2 of these passed all taken.
WHITE	2,133	1,324 or 62 %	396 or 19%	256 or 12%	157 or 7%	25 took only 1 or 2 tests. 14 of these passed all taken.
OTHER (Asian-American, Native-American, etc.)	43	20 or 47%	10 or 23%	8 or 19%	5 or 11%	3 took only 1 or 2 tests. 1 of these passed all taken.

Source: Texas Education Agency, Austin, Texas

Appendix N

JULY 1984

PERFORMANCE BY NUMBER OF TESTS PASSED

Students Who Took 3 Tests

State Totals	Number Tested	Passed Three Tests	Passed Two Tests	Passed One Test	Passed None	Note
	4217	2207 52%	834 20%	591 14%	585 14%	

HISPANIC	555	102 19%	133 24%	136 25%	184 33%	
BLACK	289	43 15%	44 15%	60 21%	142 49%	
WHITE	3283	2025 62%	643 20%	379 12%	236 7%	
OTHER (Asian-American, Native-American, etc.)	90	37 41%	14 16%	16 16%	23 26%	

Source: Texas Education Agency, Austin, Texas.

Appendix O

NOVEMBER 1984

PERFORMANCE BY NUMBER OF TESTS PASSED

Students Taking Three Tests

State Total	Number tested	Passed Three Tests	Passed Two Tests	Passed One Test	Passed None	Note:
6,209	5,123 83%	3,110 61%	964 19%	579 11%	470 9%	

Hispanic	531 10%	153 29%	122 23%	114 21%	142 27%	
Black	199 4%	39 20%	30 15%	48 24%	82 41%	
White	4,280 84%	2,852 67%	791 19%	403 9%	234 5%	
Other	113 2%	66 58%	21 19%	14 12%	12 11%	

126

Source: Texas Education Agency, Austin, Texas.

Appendix P

MARCH 1985  
 PERFORMANCE BY NUMBER OF TESTS PASSED  
 STUDENTS TAKING THREE TESTS

State Totals	Number Tested	Passed Three Tests	Passed Two Tests	Passed One Test	Passed None	Note
7,653	5,926 77%	3,554 60%	1,118 19%	656 11%	598 10%	
Hispanic	708 12%	228 32%	163 23%	156 22%	161 23%	
Black	316 5%	69 22%	52 16%	67 21%	128 41%	
White	4,772 81%	3,202 67%	873 18%	406 9%	291 6%	
Other	130 2%	55 42%	30 23%	27 21%	18 14%	

Source: Texas Education Agency, Austin, Texas.

BEST COPY AVAILABLE

NUMBER OF CANDIDATES TESTED  
1985

BEZI COLA VARIANTE

	March	
	#	%
Candidates Tested	7653	
Male	1626	21.2%
Female	6022	78.7%
Unknown	5	0.1%
Hispanic	1015	13.3%
Black	424	5.5%
White	6052	79.1%
Other	162	2.1%
Freshman	340	4.4%
Sophomores	2022	26.4%
Juniors	2876	37.6%
Seniors	1295	16.9%
Graduates	1031	13.5%
Failed to Indicate	89	1.2%
Reported G.P.A.*		
3.5-4.0	932	12.2%
3.0-3.49	2077	27.1%
2.5-2.99	2820	36.8%
2.0-2.49	1572	20.5%
1.5-1.99	190	2.5%
below 1.5	62	0.8%

\*G.P.A. refers to "grade point average." H is the average of course grades using an A as 4 points, a B as 3 points, a C as 2 points, and a D as 1 point.

Appendix R

Pre-Professional Skills Test  
 Longitudinal Data  
 March 1984 Through March 1985

<u>Candidates Tested</u>	<u>Number Taking</u>	<u>Number Passing</u>	<u>Percent Passing</u>	<u>Number not Retaking*</u>
Total	18,047	11,808	65.4	N/A
Total taking PPST once	15,101	10,439	69.1	N/A
Total taking PPST more than once*	2,946	1,369	46.5	2,451
Total taking one or more times by ethnicity:				
Black	895	203	22.7	319
Hispanic	2,136	732	34.3	631
Other	394	193	49.0	91
White	14,622	10,680	73.0	1,410

\*These values exclude examinees from March 1985 since data on retakers are not available for that group.

Source: Texas Education Agency, Austin, Texas.

### Author Notes

Preparation of this article was supported by the National Institute of Education Grant NIE-G-85-0004 to Peter A. Garcia.

The writer wishes to acknowledge the valuable contributions of others in the completion of this study. The constant support in the collection and review of data by Sylvia Garcia, the writer's wife, and Marcia Copps Smith is greatly appreciated. Gratitude is extended for the accuracy in interpretation of historical, legal and psychometrical information to Dr. Ann Jarvella Wilson, Carroll College; and Dr. Paul Swank, University of Houston. I am also thankful to the following individuals for final review of the study: Donna Gollnick, American Association of Colleges for Teacher Education; and Dr. Daniel Ulibarri, Dr. Alex Stein, and Dr. Betsy Hadden, National Clearinghouse for Bilingual Education. I also wish to express my appreciation to the Project Officer involved in this study from the National Institute of Education.

A conscientious effort has been made to interpret the voluminous research and data as objectively as possible. The author welcomes suggestions, comments, and recommendations in this work.