

Spring 1989

Legal Challenges to Testing for Teacher Certification: History, Impact, and Future Trends

Judith M. Cohen

Follow this and additional works at: <https://scholarcommons.sc.edu/jled>



Part of the [Law Commons](#)

Recommended Citation

Judith H. Cohen, Legal Challenges to Testing for Teacher Certification: History, Impact and Future Trends, 18 J.L. & EDUC. 229 (1989).

This Article is brought to you by the Law Reviews and Journals at Scholar Commons. It has been accepted for inclusion in The Journal of Law and Education by an authorized editor of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.

Legal Challenges to Testing for Teacher Certification: History, Impact and Future Trends

JUDITH H. COHEN*

1. Introduction

The quality of American education is of great national concern. Recent reports have documented that the elementary and secondary public schools are not producing well educated students, and in fact, that some graduating students lack even the most fundamental literacy skills needed to become productive citizens.¹ Various factors have been examined to determine the cause of declining quality in education, ultimately focusing upon the competency of those who teach.

Two prominent task forces in education have recommended that teacher education standards and curriculum be revised.² These reports have described curriculum reforms at the college level for preservice teacher education designed to ensure that classroom teachers will be both broadly and deeply educated in the liberal arts. The assumption is that if individuals are well educated, they will be better teachers. As yet, these suggestions have not been fully adopted, implemented or evaluated, but preservice teacher education programs have begun to strengthen the liberal arts requirements for their students.

State legislatures have responded to the issue of teacher competency by intervening in the certification process.³ Those wishing to qualify for

* Associate Professor of Education, Adelphi Univ. B.A. 1966 Queens College, C.U.N.Y.; M.S. 1968 Syracuse Univ.; Ph.D. 1975 & J.D. 1988 Hofstra Univ. The author gratefully acknowledges the cooperation of several people who contributed information to this investigation. In particular Michael Rebell, (Rebell & Katzive, Attorneys, New York, New York) and staff, Scott Elliott, Ed. D. (National Evaluation Systems, Amherst, Mass.) as well as staff in the state education agencies of New York, Texas, and Alabama facilitated the work by providing information and documents and documents not easily available to the general public.

1. See, e.g., NATIONAL COMMISSION ON EXCELLENCE IN EDUCATION, *A NATION AT RISK: THE IMPERATIVE FOR EDUCATIONAL REFORM* (1983); E. BOYER, *HIGH SCHOOL: A REPORT ON SECONDARY EDUCATION IN AMERICA* (1983); J. GOODLAD, *A PLACE CALLED SCHOOL: PROSPECTS FOR THE FUTURE* (1984); T. SIZER, *A REVIEW AND COMMENT ON THE NATIONAL REPORTS* (1983).

2. See CARNEGIE FORUM ON EDUCATION AND THE ECONOMY, *A NATION PREPARED: TEACHERS FOR THE 21ST CENTURY* (1986); THE HOLMES GROUP, *TOMORROW'S TEACHERS* (1986).

3. An example of recent state action in teacher certification can be found in New York. In 1980 the Regulations of the Commissioner of Education were amended to become effective on September

teacher certification are now often required to pass an objective examination in addition to completing an approved course of study. In prior times, candidates had only to demonstrate that they completed approved courses of study in accredited colleges and were automatically granted initial state certification without further evaluation. It is now possible for candidates to complete an approved course of study in a registered college program, be recommended for certification by a college, and then not qualify because of failure on an objective test. Individuals in this situation have legally challenged the validity of the examinations and the imposition of what is seen to be arbitrary cutoff scores.⁴

The use of standardized, objective teacher examinations has become widespread. As of 1984, 38 states use a test as part of certification procedures⁵ and by 1988 nine more states are expected to require competency tests.⁶ The increased use of these examinations has resulted in a growing amount of litigation. This is particularly true for minority candidates, because of the highly adverse impact these tests have on them.⁷

The competency testing movement grows out of a general perception that teachers should be held accountable for the quality of education. The public has demanded assurance that only qualified teachers educate today's youth. More than 85% of people surveyed in recent national polls feel that teachers should be required to pass exams in the subject areas they teach.⁸ Lack of teacher competence is often used to explain declining student achievement over the past 15 years.⁹ While long term measures to raise the quality of those desiring to become teachers have been con-

2, 1984. The new regulations required that all applicants for their first New York State certificate had to file an official score report from Educational Testing Service confirming the attainment of set scores on the NTE Core Battery Tests. Subsequently, the regulations were again amended so that all candidates requesting initial certification would need to submit evidence of having taken a Specialty Area Test. This was to become effective in September 1987. No passing levels for the specialty tests were determined, but a candidate was to submit evidence of taking a specialty test for each teaching area in which certification is sought. Subsequently, the imposition of this regulation regarding specialty tests was frozen by the new Commissioner of Education.

4. A parallel movement to adopt competency tests for inservice teachers (practicing teachers are usually referred to as "inservice teachers," whereas those preparing to become teachers are referred to a "preservice") appears to be "grinding to a halt." According to the National Education association "no state has enacted laws requiring competency tests for practicing teachers since 1985. Only Arkansas, Georgia, and Texas now have such tests." *Testing*, Educ. Week, May 20, 1987, at 4, col. 5.

But see, Note, *Teacher, Termination and Competency Testing*, 63 TEX. L. REV. (1985) (arguing for the imposition of competency testing for certified teachers).

5. W. Gorth & M. Chernoff, *TESTING FOR TEACHER CERTIFICATION I* (1986).

6. Rebell, *Disparate Impact of Teacher Competency Testing on Minorities: Don't Blame the Test-Takers — or the Tests*, 4 YALE L. & POL'Y REV. 375 (1986).

7. *Id.* at 376.

8. Vorwerk & Gorth, *Common Themes in Teacher Certification Testing, Program Development and Implementation*, in *TESTING FOR TEACHER CERTIFICATION* 36 (1986).

9. *Id.*

sidered, the public appears to be demanding immediate action to protect children from what is often perceived as incompetent teachers. "In general, as accountability in education becomes more important, the impetus for testing prospective teachers grows."¹⁰

The use of standardized tests to measure the competency of prospective teachers appears to assure that the individual to be certified is qualified to teach. The appeal of objective testing is both theoretical and practical. By design, objective tests provide an efficient and inexpensive way to measure how well an individual has mastered a set of external, knowledge-based competencies. Presently, states that use a teacher certification testing program have three options for obtaining such tests: custom developed tests, off-the-shelf tests, and modified customized tests.¹¹ The National Teacher Exam ("NTE"), consisting of the Core Battery Tests and Specialty Area Tests, is the most widely used objective test for teacher certification.¹²

However, the wide use of these standardized tests raises many significant concerns. It has been noted that there is a disparate impact between the passing rates of minority group members as compared to white candidates and therefore there are likely to be fewer minority candidates eligible to teach.¹³ No one would desire that poorly educated, minority individuals become teachers and perpetuate inferior education. But evaluation systems that negatively affect the availability of significant numbers of minority candidates to qualify as teachers, is of great concern. "That thousands of black and Hispanic students who have prepared for teaching careers are being denied entrance certification to a profession in which minorities are already under represented is a serious and pressing social issue."¹⁴

An even broader question is whether or not paper and pencil tests are able to accurately predict and measure the skills required for effective teaching. While good teaching is undeniably associated with a knowledge

10. Gorth & Chernoff, *supra* note 5, at 3.

11. *Id.* at 10. The development of custom tests involves creating a set of original test questions keyed to objectives. The questions are then validated through a job analysis survey. Off-the-shelf tests are standardized, objective tests administered according to the established test administration policies with no modifications allowed. A modified customized test uses pre-existing test items specifically adapted to the specific testing situation.

12. The Core Battery Tests of the National Teachers Examination ("NTE") consists of three separate tests: Communication Skills (reading, writing and listening), General Knowledge (liberal arts areas of literature, fine arts, mathematics, science and social studies), and Professional Knowledge (social and cultural principles that influence curriculum and general principles of learning and instruction.) The Specialty Area Tests of the NTE consists of tests designed to measure the content and methods applicable to teaching specific subject areas.

13. Rubinstein, *The Changing Nature of Teacher Certification Programs*, in TESTING FOR TEACHER CERTIFICATION 31 (1986).

14. Rebell, *supra* note 6, at 376-77.

base, much of teaching is dynamic and interactive and does not lend itself to measurement by use of multiple choice tests. It is recognized that current teacher tests "leave out an awful lot that is associated with good teaching . . . and used alone such tests have tended to trivialize and reduce teaching to a few measurable, but 'scientifically validated,' kinds of skills."¹⁵

This article surveys the legal challenges to certification based on the use of objective competency tests and analyzes the theories under which such challenges have been brought. Furthermore, recent trends created by the implementation of these tests will be discussed with particular reference to the declining number of eligible minority candidates. Finally, it will describe alternative models to off-the-shelf, competency evaluations by the creation of "custom designed" certification programs presently used by a few states. Another very recent proposal to create national standards for professional certification of teachers by a peer review system will also be described. It is hoped that the design of a new series of performance based tests will revolutionize the assessment of teachers and "should sound the death knell for many tests now in use."¹⁶

If society values a diverse and fully integrated teaching staff, one with demonstrated ability to perform effectively, then other forms of competency evaluation for certification will continue to be developed and utilized. It is anticipated that the new forms of evaluation will withstand legal challenges and contribute to the assurance that certified candidates are truly competent to teach.

2. Statutory Regulation of Teacher Certification

Every state has enacted statutes regulating standards for becoming a public school teacher.¹⁷ "Certification is the process of legal sanction which authorizes the one certified to perform specific services in the public schools of the state."¹⁸ The state agency that grants certification determines the qualifications for different teaching licenses, and while certification standards vary from state to state, they most often contain one or more of these elements:

- completion of postsecondary course of study

15. *Stanford Project Seeks Better Ways to Assess Teacher Performance*, Educ. Week, April 8, 1987 at 18, col. 3 (quoting Gary Sykes, Associate Director of Stanford research project).

16. Hechinger, *Testing and Teachers*, N.Y. Times, May 19, 1987, at C12, col. 4.

17. Citron, *An Overview of Legal Issues in Teacher Quality*, 14 J.L. & Educ. 277, 278 (1985).

18. Note, *Minimum Competency Testing of Teachers for Certification: Due Process, Equal Protection and Title VII Implications*, 70 Cornell L. Rev. 494 (1985) (citing Gardner & Palmer, *Certification and Accreditation: Background, Issue Analysis, and Recommendations 3-5*) (Report Prepared for the National Commission on Excellence in Education, Aug. 1982).

- attainment of a fixed grade point average in collegiate studies
- completion of an internship experience commonly referred to as student teaching
- passing a competency test or tests often composed of basic skills sections (communication skills, background knowledge of the liberal arts, and general knowledge about education), and special area sections (specific knowledge of a chosen teaching area).¹⁹

Law that governs certification serves several purposes. "It fosters adequate professional training to ensure competence in the licensed area; maintains registration of teachers with state and district agencies to ensure compliance with certification law and enforce certification guidelines; and protects certified teachers to some degree against unfair displacement."²⁰ The certification process is within the legitimate scope of state interest in creating quality education and has not been successfully challenged absent an error of law or arbitrary action.²¹ When employment decisions are based primarily on academic considerations, courts give deference and broad discretion to state agencies.²²

In the 1970's, with the heightened concern for quality education, many states modified certification standards to insure a more strenuous demonstration of teacher competency. Often this resulted in the inclusion of standardized teacher certification tests. "The State's authority to redefine minimal competence from time to time cannot reasonably be questioned."²³ However, when minimal competency required a passing score on an objective examination, unsuccessful candidates brought their complaint to the courthouse. "Testing teachers has been perhaps the most controversial screening strategy and consequently the most litigated."²⁴ In the leading case in this area, *United States v. South Carolina*, the court upheld a state's right to impose a competency test on prospective teachers as furthering an important governmental objective:

The evidence in the record supports a finding that South Carolina officials were concerned with improving the quality of public school teaching, certifying only those applicants possessed of the minimum knowledge necessary to teach effectively, utilizing an objective measure of applicants coming from widely disparate teacher training programs, and providing appropriate financial incentives for teachers to improve their academic qualifications and thereby their ability to

19. Citron, *supra* note 17, at 278 & 281.

20. W. VALENTE, *LAW IN THE SCHOOLS* 181 (1987).

21. *Id.*

22. Swiger & Zehr, *The Search for Excellence: Legal Issues in Teacher Competency Testing*, 16 *URB. LAW.* 745, 748 (1984).

23. *United States v. South Carolina*, 445 F. Supp. 1094, 1103 (D.S.C. 1977), *aff'd*, 434 U.S. 1026 (1978).

24. Citron, *supra* note 17 at 280.

teach. We conclude that these are entirely legitimate and clearly important governmental objectives.²⁵

3. Title VII and Certification Testing

A. *A New Cause of Action for Plaintiffs*

Those challenging the use of a standardized test for certification have brought claims on a variety of theories including constitutional grounds citing violations in both due process and equal protection.²⁶ When the 1972 Civil Rights Amendment extended Title VII of the Civil Rights Act of 1964 to state and local government employers, minority teachers were given an additional basis on which to challenge state practices in employment that resulted in a disparate impact.

In the first round of successful challenges to testing for teacher certification in the 1970's, plaintiffs prevailed under Title VII of the Civil Rights Act of 1964 which prohibits employment discrimination based on sex, race, color, religion or national origin.²⁷ The Equal Employment Opportunity Commission (EEOC) was empowered to enforce this law. Title VII has been applied in a variety of employment contexts that include both overt discriminatory practices as well as employment practices that result in the elimination of minority representation. Thus tests that resulted in the exclusion of minority individuals from employment as teachers came under the scrutiny of Title VII standards. If the use of an employment test resulted in a disparate impact for minority individuals, an initial inquiry was made to determine if it constituted a significant, disparate impact. If such disparate impact was determined, defendants could justify the use of the practice in question if it could be shown to be job-related. In the case of teacher testing, defendants could meet the job-related standard by showing that the instrument in use was developed under appropriate professional standards. "In this way, the suspicion that an ostensibly objective test was being used as a 'pretext' for a hidden discriminatory motive will have been set aside."²⁸ The publishing of specific guidelines by professional groups in education and psychology created objective standards for testing that assisted in the determination of the fairness of these certification tests. "The promulgation of these regulations and standards

25. *South Carolina*, 445 F. Supp. at 1108.

26. U.S. Const. amend. XIV, § 1.

27. Title VII of the Civil Rights Act of 1964, 42 U.S.C. §§ 2000e to 2000e-15 (1982) [hereinafter Title VII].

28. Rebell, *Legal Issues Concerning Bias in Testing*, in *BIAS ISSUES IN TEACHER CERTIFICATION TESTING 2* (1988).

reflected increasing legislative, judicial, and professional concern with fair employment practices both in and out of education."²⁹

In 1970 EEOC guidelines, that revised the original 1966 guidelines, included criteria for determining if employee selection procedures were non-discriminatory as required by Title VII. The guidelines included the following provisions that apply to evaluating the fairness of employment testing in such areas as certification:

- Empirical data should be made available to establish the predictive validity of a test, that is, the correlation of test performance with job-relevant work behaviors; such data should be collected according to generally accepted procedures for establishing criterion-related validity.
- Where predictive validity is not feasible, evidence of content validity (in the case of job knowledge or proficiency tests) may suffice as long as appropriate information relating test content to job requirements is supplied.
- Where validity cannot otherwise be established, evidence of a test's validity can be claimed on the basis of validation in other organizations as long as the jobs are shown to be comparable and there are no major differences in context or sample composition.
- Differential failure rates (with consequent adverse effects on hiring) for members of groups protected by Title VII constitute discrimination unless the test has been proven valid (as defined above) and alternative procedures for selection are not available.
- Differential failure rates must have a job-relevant basis and, where possible, data on such rates must be reported separately for minority and nonminority groups.³⁰

These guidelines require that Title VII challenges to certification testing deal with the issue of test validity. Test validity is "the degree to which the test actually measures what it purports to measure. Validity provides a direct check on how well the test fulfills its function, [and] the determination of validity usually requires independent, external criteria of whatever the test is designed to measure."³¹ The intent of the EEOC guidelines is specific: "A test must be a representative measure of the actual domain of skills used on the job and must be validated for its intended purpose."³²

B. Title VII Standards for Employment Tests

A major challenge to the use of an employment test, resulting in adverse minority impact, was decided by the Supreme Court in 1971. In *Griggs v.*

29. Rubenstein, *supra* note 13 at 26.

30. *Id.* at 24-25.

31. ANASTASI, *PSYCHOLOGICAL TESTING* 27 (1982).

32. Rubenstein, *supra* note 13 at 26.

*Duke Power Co.*³³ thirteen black petitioners sought to overturn the use of a standardized intelligence test as a condition for employment or job transfer. Although the North Carolina power company had a policy of overt discrimination prior to the enactment of Title VII, such conduct had stopped.³⁴ The lower court found that there was no racial discrimination or other invidious intent in the imposition of the test, but rather that the company truly intended to improve the quality of the work force.³⁵ However, the effect of imposing the test was "to render ineligible a markedly disproportionate number of Negroes."³⁶

Reviewing the purpose of Title VII the *Griggs*' court concluded that "The Act proscribes not only overt discrimination but also practices that are fair in form, but discriminatory in operation."³⁷ Furthermore, "If an employment practice which operates to exclude Negroes cannot be shown to be related to job performance, the practice is prohibited."³⁸ Because the test was not shown to bear a demonstrable relationship to successful job performance, its use was prohibited.

Griggs has come to stand for the "removal of artificial, arbitrary, and unnecessary barriers to employment when the barriers operate invidiously to discriminate on the basis of racial or other impermissible classification."³⁹ The practice of proving job relatedness, whenever employment testing is imposed, thus became an essential concept. "Broad and general testing devices . . . as fixed measures of capability" would be barred in employment testing.⁴⁰

Job relatedness in testing was further clarified in *Albermarle Paper Co. v. Moody*.⁴¹ Again, black employees challenged employment testing that prevented their promotion to higher paying jobs. Here, the court disfavored validation studies conducted for the employment tests as materially defective and emphasized that to determine if a test was job relevant, one needed to evaluate a particular skill to perform a specific job.

4. Survey of Legal Challenges to Certification Testing

A. Southern Civil Rights Cases from the 1970's

In addition to the South Carolina case, the 1970's saw a significant

33. 401 U.S. 424 (1971).

34. *Id.* at 428.

35. *Id.* at 431.

36. *Id.* at 429.

37. *Id.* at 431.

38. *Id.*

39. *Id.*

40. *Id.* at 433.

41. 422 U.S. 405 (1975).

number of cases litigated in the states of Georgia, North Carolina, Mississippi, and Virginia.⁴² These cases were brought by minority plaintiffs who challenged the use of a passing grade on the National Teacher Examination ("NTE") as a certification requirement. The cases presented these common elements: a history of prior school desegregation, the sudden imposition of a certification test that was not validated for this purpose, use of arbitrary cut-off scores, and large numbers of black examinees who were often poorly educated in racially segregated schools. In all of the cases, use of the NTE clearly resulted in adverse, disparate impact for blacks, who were disqualified from teaching employment.⁴³

When use of the NTE adversely affected the ability of minority candidates to qualify for teaching licenses, the litigants sought relief by citing violations of both Title VII and the fourteenth amendment. An essential factor in the determination of these cases was the lack of job-related validity for the NTE as a certification test. This certainly was a clear violation of Title VII employment testing guidelines.

At the time these cases were litigated, Educational Testing Service ("ETS"), publisher of the test, did not recommend the use of the test for granting certification, because it had not been created for that purpose. "Consistently, Educational Testing Service had urged the state Board *not* to use NTE scores as cut-off requirements for licensing."⁴⁴ Following the Supreme Court's decision in *Griggs*⁴⁵ the courts held that employment tests must have job-related validity. No evidence could be brought that correlated a score on the NTE with teacher competency. "It was uncontradicted that the NTE cut-off score requirement was set without any investigation or study of the validity and reliability of the examination or the cut-off score as a means of selecting teachers for hiring or re-employment, and without consultation with the Educational Testing Service."⁴⁶ The NTE was measuring the person in the abstract and not for the specific job skills.⁴⁷ There was no "*demonstrable relationship* to successful performance of the jobs for which it was used."⁴⁸

Because testing with the NTE resulted in stratifying teachers by race, its

42. See, e.g., *Baker v. Columbus Mun. Separate Sch. Dist.*, 329 F. Supp. 706 (N.D. Miss. 1971), *aff'd*, 462 F.2d 1112 (5th Cir. 1972); *Armstead v. Starkville Mun. Separate Sch. Dist.*, 461 F.2d 276 (5th Cir. 1972); *Walston v. County Sch. Bd. of Nansemond County*, 492 F.2d 919 (4th Cir. 1973); *United States v. North Carolina*, 400 F. Supp. 343 (E.D.N.C. 1975), *vacated*, 425 F. Supp. 789 (E.D.N.C. 1977); *Georgia Assoc. of Educators v. Nix*, 407 F. Supp. 112 (N.D. Ga. 1976).

43. In *Walston*, whereas 38% of the teachers hired in 1970-71 were black, the number declined to 14% in 1971-72 when the examination was imposed. *Walston*, 492 F.2d at 922.

44. *United States v. North Carolina*, 400 F. Supp. at 348.

45. *Griggs*, 401 U.S. 424 (1971).

46. *Baker*, 462 F.2d at 1114.

47. *Walston*, 492 F.2d at 926.

48. *Id.* at 924 (citing *Griggs*, 401 U.S. at 431).

use was subjected to equal protection analysis. The strict scrutiny standard of review is applied whenever classification is based on a suspect category such as race. For this employment test to withstand the equal protection attack, it had to be justified by an overriding purpose independent of a racial effect.⁴⁹ The Baker court concluded:

We find no such overriding purpose here. While a school board's desire to improve the caliber of its faculty is a laudable one, the policies and procedure employed must be clearly and fairly related to this goal. The totality of the record, on the contrary, indicates a strong undercurrent of discrimination that has effectively decimated the ranks of black teachers whose credentials are equally, if not more impressive than many of those hired.⁵⁰

The courts concluded that the use of the NTE, as a sole criterion for teacher employment, should be prohibited. They balanced the legitimate interest of school boards that claimed to be concerned with quality education with the desire that school boards make certain that "teachers are not the victims of arbitrary, discriminatory action."⁵¹ The courts were especially sensitive to the plight of black teachers, many poorly educated themselves in segregated, southern schools, who had the potential for becoming effective teachers. Denying employment by the imposition of a invalid examination seemed particularly unfair and would perpetuate vestiges of racial segregation. In summary, these courts had no difficulty prohibiting the use of the NTE, which was not validated for certification purposes, in school systems that were recently desegregated and that resulted in prohibiting employment for significant numbers of black teachers.

B. *The Impact of Washington v. Davis*

In its 1974 opinion the court in *Walston* foreshadowed future litigation when it said; "[i]t may well be that when properly and fairly applied in appropriate situations, the NTE could qualify as having the 'demonstrable relationship' as required by *Griggs* so essential to ensure equal employment opportunity."⁵² Examination of the two hearings of *United States v. North Carolina*⁵³ in 1975 and in 1977 demonstrates that the use of a standardized test, even one that results in disparate impact, will be upheld as a

49. *Baker*, 462 F.2d at 1114.

50. *Walston*, 492 F.2d at 925-26.

51. *See id.* at 927.

52. *Id.*

53. *United States v. North Carolina*, 400 F. Supp. 343 (E.D.N.C. 1975), *vacated*, 425 F. Supp. 789 (E.D.N.C. 1977). [hereinafter *North Carolina*]

requirement for teacher certification if it is substantially related to job performance.

In the 1975 hearing of the *North Carolina* suit, the United States brought a Title VII complaint for "invidious discrimination against blacks, Indians and Oriental persons implemented by minimum test score requirements on the NTE."⁵⁴ The discriminatory impact of the NTE was startling. The failure rate for black examinees was 31.08%, whereas it was 1.36% for white candidates.⁵⁵ The court was sympathetic to the dilemma of the state certifying agency that was unable to rely solely on certification recommendations from the 32 different teacher training institutions within the state. There was evidence of marked variations in the quality and standards of teacher education programs throughout the state. Equally troubling, however, was the plight of black candidates, poorly educated in segregated public schools, who then entered inferior teacher education programs, most often in predominantly black colleges. The record revealed that the president of one teacher training institute admitted to having sometimes graduated functional illiterates and the court acknowledged that the state should have "the right to adopt academic requirements and written achievement tests designed and validated to disclose the minimum amount of knowledge necessary to effective teaching."⁵⁶ However, the NTE was not designed for use in assessing inservice teachers, the cut-off score chosen was not validated for job performance, and the result was a disparate impact on blacks. The court held that the use of the NTE "constitute[d] a burdensome and arbitrary denial of equal protection under the fourteenth amendment."⁵⁷

To reach its conclusion, the court balanced the right of access by people in a free society to all vocations with the state's interest in protecting the public from incompetent teachers. Ultimately, "[t]he right of the State to set standards for the purpose of improving the quality of instruction in the public schools [was] not separable from the right of the prospective teacher applicant to enter his chosen vocation."⁵⁸ The use of the NTE was still prohibited, because it had not been validated to demonstrate the necessary job-relationship. Shortly thereafter this decision was vacated because of the Supreme Court's ruling in *Washington v. Davis*⁵⁹ and a validation study conducted for the NTE in North Carolina.

54. See *id.* at 346.

55. See *id.* at 347.

56. See *id.* at 348.

57. See *id.* at 349.

58. See *id.* at 350.

59. 426 U.S. 229 (1976).

In *Washington v. Davis*⁶⁰ the Supreme Court reviewed the use of a verbal skills test for entry into a police training program. Use of the test was challenged because it was contended not to have any relationship to job performance and because it excluded a disproportionately high number of black applicants. Here, the Supreme Court did not recognize that "racially differential impact" was by itself sufficient proof of a racially discriminatory purpose. An "invidious discriminatory purpose [might] often [have] been inferred from the totality of the relevant facts, including the fact, if it [was] true, that the law bears more heavily on one race than another."⁶¹ But the Court had "difficulty understanding how a law establishing a racially neutral qualification for employment [was] nevertheless racially discriminatory and denie[d] 'any person . . . equal protection of the law' simply because a greater proportion of Negroes fail[ed] to qualify than members of other racial or ethnic groups."⁶² Additionally, the use of the test was supported by its positive correlation with training-course performance as demonstrated by a validation study. Title VII standards were not applied in *Washington v. Davis* because the statute was not applicable to federal employees when the case was initially filed.⁶³ In the *Washington v. Davis* analysis the Court clearly stated that Title VII standards were not identical to the equal protection standards where a discriminatory purpose had to be demonstrated to render a practice unconstitutional.⁶⁴ However, the Court did comment that if the job-relatedness requirement of *Griggs* or *Albermarle Paper* were to be applied, the evidence here supported a "positive relationship between the test and training-course performance . . . sufficient to validate the former, wholly aside from its possible relationship to actual performance as a police officer."⁶⁵ Thus the use of a test which appeared "neutral on its face and rational may be said to serve a purpose the Government is constitutionally empowered to pursue" was allowed under equal protection analysis even though its use resulted in exclusion of a significant number of minority applicants.⁶⁶

Applying this new analysis, the 1975 decision in *North Carolina* was vacated in 1977. While "disproportionate impact is not irrelevant . . . it is not the sole touchstone of an invidious racial discrimination."⁶⁷ "Disproportionate racial impact of teacher employment tests does not

60. *Id.*

61. *See id.* at 242.

62. *See id.* at 245.

63. *See id.* at 238.

64. *See id.* at 239.

65. *See id.* at 250.

66. *See id.* at 246.

67. *North Carolina*, 425 F. Supp. at 793.

automatically establish unconstitutional discrimination," and no longer was there a single method of appropriately validating employment tests for job-related performance.⁶⁸

The new analysis was applied to a teacher certification context in 1978. In *United States v. South Carolina*⁶⁹ the NTE was challenged as violating both equal protection and Title VII. In order to sustain an allegation of a Fourteenth Amendment violation, the plaintiffs needed to prove that by using the NTE, the state intended to create and use a racial classification.⁷⁰ The court's analysis cited the historical background, the prior sequence of events, legislative history, and testimony from state officials. No intent to discriminate by race was found.⁷¹ In particular, the record indicated that the racial composition of South Carolina's teacher staff had closely paralleled the racial composition of the state population. Moreover, the NTE was chosen for certification purposes because it could be scored objectively and impartially and the state would not be accused of using it for discriminatory purposes.⁷²

To validate the particular cut-off score chosen by the state, an extensive validation study was conducted over three months resulting in a 300-page report.⁷³ The State Board adopted the report recommendations and accordingly set separate minimum score requirements for each teaching field. Use of the NTE created classifications based "only on permissible bases (presence or absence of knowledge or skill and ability in applying knowledge) . . . and their use in making certification decisions by the State is proper and legal."⁷⁴ While the NTE was not found to measure teaching skills, it was found to "measure the content of the academic preparation of prospective teachers [with] no evidence that the required passing grade was set at an arbitrarily high level."⁷⁵ Therefore, its use was upheld.

To counter the Title VII argument about disparate impact on employment, the state conducted an elaborate study to demonstrate the test's content validity. Content validity was defined as the extent to which questions on the test matched subject matter that was part of the curriculum at the state's teacher training institutions.⁷⁶ The review panel, consisting of

68. Valente, *supra* note 20 at 367.

69. *South Carolina*, 445 F. Supp. 1094 (D.S.C. 1977), *aff'd*, 434 U.S. 1026 (1978).

70. *See id.* at 1100.

71. *See id.* at 1102.

72. *See id.* at 1101.

73. *See id.* at 1104.

74. *Id.*

75. *See id.* at 1108.

76. *See id.* at 1112.

456 teacher educators, rated the questions both to their match with curriculum content and to the extent to which they had to be mastered to be a minimally qualified teacher candidate in South Carolina.⁷⁷ The validity study was found to be sufficiently trustworthy as to test content and cut-off scores.⁷⁸

An analysis of an alternate means of certifying teachers by graduation from an approved program was found insufficient for achieving the state purpose.⁷⁹ In summary, the court examined the balance between using a standardized test as part of a decision-making process with the concern for certifying only competent teachers. It upheld the use of the test to achieve the state's purpose in certifying candidates demonstrated to be at least minimally competent. The decision was affirmed in 1978 by the Supreme Court.⁸⁰ Justices White and Brennan dissented, focusing on the fact that the NTE was validated solely as to its curriculum validity and was not validated against job performance.

The decisions in *South Carolina* and *North Carolina* demonstrate that content validation techniques will suffice in lieu of the earlier "job-related" standard⁸¹ proposed by the *Griggs*' court.⁸² When teacher certification tests are professionally developed in good faith to insure teacher competency and are then validated as to content, they will be upheld by courts. The public interest in having at least minimally competent teachers seems to outweigh the disparate impact that has often resulted.

C. Challenges Based on Violation of Due Process

An alternate theory for challenging certification testing is based upon denial of due process.⁸³ Under this theory a plaintiff would need to show that he or she was deprived of a constitutionally protected liberty or property right without adequate notice or without the opportunity to be heard.

In *Board of Regents v. Roth*⁸⁴ the Supreme Court articulated a definition for a protected liberty or property interest in a teaching context. Roth was hired to teach for one academic year at a state university and was

77. *Id.*

78. *See id.* at 1114.

79. *See id.* at 1115.

80. *See South Carolina*, 445 F. Supp. 1094 (D.S.C. 1977), *aff'd*, 434 U.S. 1026 (1978).

81. The job-related standard is referred to as criterion-related validity. Criterion related validity refers to analysis of whether a test measures criterion knowledge to be possessed by effective teachers. Such tests are felt to be predictive of competence in the profession. *See Rebell, supra* note 6, at 381.

82. *See Griggs*, 401 U.S. 424 (1971).

83. The Fourteenth Amendment provides: "Nor shall any State deprive any person of life, liberty, or property without due process of law." U.S. CONST. amend. XIV, § 2.

84. *Board of Regents v. Roth*, 408 U.S. 565 (1972).

given no explanation for his nonrenewal. The Court held that due process would be violated only if Roth had a protected liberty or property interest. However, when Roth was not rehired, the state

did not base the nonrenewal of his contract on a charge, for example, that he had been guilty of dishonesty, or immorality . . . for 'where a person's good name, reputation, honor, or integrity is at stake because of what the government is doing to him, notice and an opportunity to be heard are essential.'⁸⁵

In such situations, due process would present an opportunity to refute the charges. Furthermore, the Court defined "property" as to its attributes: "To have a property interest in a benefit, a person clearly must have more than an abstract need or desire for it. He must have more than a unilateral expectation of it. He must, instead, have a legitimate claim of entitlement to it."⁸⁶ While Roth had an abstract concern in being rehired he did not have a property interest sufficient to require a hearing before the nonrenewal of his contract.⁸⁷ It is therefore unlikely that a non-tenured person who is merely an applicant for a teaching position could prevail under due process protection of a property right.⁸⁸

In an important challenge to competency tests for high school students, the court in *Debra P. v. Turlington*⁸⁹ found that not giving a high school diploma to students, who failed a minimum competency test, was a denial of due process. The district court issued a four year injunction against the use of the competency test because the test's implementation provided insufficient notice.⁹⁰ In addition, the district court found that high school students had "a liberty interest in being free of the adverse stigma associated with [failure to earn a normal diploma]."⁹¹ Even though a liberty interest was found in not obtaining a high school diploma, the denial of teacher certification causes less stigma than the denial of a high school diploma. Failure to receive the high school diploma is associated with functional illiteracy and often results in the inability to secure employment.⁹² Denial of a teaching certificate is not analogous.

There are several factors that distinguish the situation in *Debra P.* from the teacher certification challenges. High school education is compulsory,

85. See *id.* at 573 (citing *Wisconsin v. Constantineau*, 400 U.S. at 437).

86. See *id.* at 577.

87. See *id.* at 578.

88. Note, *supra* note 18, at 498.

89. *Debra P. v. Turlington*, 474 F. Supp. 244 (M.D. Fla. 1979), *aff'd in part and vacated in part*, 644 F.2d 397 (5th Cir. 1981), *on remand* 564 F. Supp. 177 (M.D. Fla. 1983), *aff'd*, 730 F.2d 1405 (11 Cir. 1984).

90. See *Debra P.*, 730 F.2d at 1407.

91. Note, *supra* note 18, at 500.

92. *Id.*

whereas collegiate level teacher education is not. The high school diploma is the result of 12 years of schooling, while teacher education programs are usually only 4 years in duration. Finally, the inability to obtain teacher certification does not preclude other employment opportunities, while failure to obtain a high school diploma prevents many employment opportunities.⁹³ Additionally, much of the court's analysis in *Debra P.* was based on the high school student's "legitimate expectation" in obtaining a diploma if he or she completes the course of study. Teaching candidates do not have the same legitimate expectation of obtaining the credential as do high school students. Obtaining state certification rests on the concept of being properly well qualified. Because the public should be protected from incompetent teachers, merely attending college for four years does not automatically entitle a candidate to certification.⁹⁴ It is therefore unlikely that the competency testing situation in *Debra P.* could be used to support a due process argument in certification testing.

The denial of due process was presented as an argument in *United States v. Lulac*⁹⁵ where minority candidates challenged Texas' imposition of the Pre-Professional Skills Test ("PPST") developed by ETS as an entrance requirement for admission into advanced teacher education course. The court agreed with the plaintiffs that "[a]s a matter of substantive due process . . . an individual has a liberty right to engage in a chosen profession free from unreasonable governmental interference."⁹⁶ However, the court further reasoned that the corollary to this proposition would be that a reasonable governmental examination would be permitted. If a test is found to be valid, it would not be demonstrably unreasonable and the due process argument would fail as it did here.

If a state modifies its teacher certification requirements with adequate notice to potential candidates, it is unlikely that it will be successfully challenged based on the theory of denial of a property interest. The *Lulac* court held that the legislative process, itself, gives adequate notice. "When the legislature enacts a law, or a state agency adopts a regulation, that affects a general class of persons, all of those persons have received procedural due process by the legislative process itself and they have no right to individual attention."⁹⁷ However, to avoid the problem of inadequate notice, many states have allowed several years to elapse before new standards are imposed. Other states publish test results from the teacher

93. See *id.* at 502.

94. Rebell, *supra* note 6, at 390.

95. *United States v. Texas*, 628 F. Supp. 304 (E.D. Tex. 1985), *rev'd sub nom.* *United States v. Lulac*, 793 F.2d 636 (5th Cir. 1986).

96. *Lulac*, 793 F.2d, at 647.

97. See *id.* at 648.

education institutions to help candidates choose programs where they can expect to be well prepared to pass certification exams.⁹⁸

D. Challenges Based on Lack of Test Validity

A complex argument that is often the focus of current challenges to certification testing is based on the issue of test validity. The Supreme Court has not reviewed the concept of test validity since its major rulings ten years ago. Therefore, the lower courts have attempted to apply the concept of test validation recently in the context of high school competency testing.⁹⁹

As part of the movement to raise standards in education, most states have enacted minimum competency tests for receipt of a high school diploma. High school competency tests usually evaluate a student's mastery of basic skills in reading, writing, and mathematics. Analogous to the imposition of the NTE, these minimum competency tests have had disparate impact on minority students. At the time of the first hearing of *Debra P.* in 1979 and after three administrations of the competency tests in Florida, the failure rate of black students was ten times greater than that of white students.¹⁰⁰ A four year injunction was imposed against the use of a test for denying high school diplomas. Subsequently, the Court of Appeals focused its inquiry into the reimposition of the competency tests by determining if the tests were instructionally valid.

Instructional validity is the attempt to determine if students are given the opportunity to learn what is tested on the competency exams, but "the *Debra P.* remand hearing shows that any attempt to determine what individual students have actually been taught is impossible."¹⁰¹ To demonstrate actual instructional validity, one would need to show how well each individual child was taught. This would establish that each individual had a fair opportunity to learn the subjects on the examination.

To attempt an instructional validation of Florida's competency tests an elaborate study was conducted. An independent assessment firm was hired to design and implement a study that consisted of several parts: (1) teacher survey distributed to all of Florida's 65,000 teachers (47,000 teachers responded to the survey) to determine if they provided instruction in basic skills; (2) district survey of all of Florida's school districts to determine at which grade levels competency skills were taught and expected to be mastered; (3) series of site visits to school districts to verify the accuracy of

98. Note, *supra* note 18, at 502-03.

99. Rebell, *supra* note 6, at 384.

100. See *Debra P.*, 730 F.2d at 1406.

101. Rebell, *supra* note 6, at 387.

the reports in (2); and (4) survey of 3,200 randomly selected students attending school visited to determine whether they had been taught skills tested by the competency exams.

In spite of the ambitious scope of this study, the district court candidly admitted its frustration at determining instructional validity, citing the impossibility of measuring each child's opportunity for learning.¹⁰² Instead, the district court substituted a less demanding test of validity, called curricular validity, because "there are no accepted education standards for determining whether a test is instructionally valid."¹⁰³ Curricular validity was defined as an attempt to determine if the skills tested were included in the official school curricula and that the majority of teachers recognized the curricular goals as being appropriate teaching goals.¹⁰⁴ Ultimately, Florida's competency testing program was upheld because it met the standard for curricular validity. In fact, the court was positively impressed with the impact the competency program had on school reform. The imposition of the testing program created teaching reforms, including extensive remedial programs, and thereby removed past vestiges of inferior schooling for black students.

The court in *Lulac*¹⁰⁵ has also very recently dealt with the concept of test validity and teacher certification. In 1981 the Texas legislature adopted a law requiring satisfactory performance on a basic skills test, the PPST, as a condition of admission into an approved teacher education program. The test was designed to measure reading, writing, and arithmetic skills at a 12th grade difficulty. A district court issued a preliminary injunction against the use of the test, but it was reversed for not considering evidence of the test's validity.

The challenge to imposition of the PPST was felt to "turn on whether the test [was] a fair measure of necessary skills."¹⁰⁶ The same consulting firm that conducted the validation study in *Debra P.* was hired to conduct a similar study in Texas. A lengthy validation report based on the assessment of 624 Texas university and public school teachers demonstrated that more than 95% of the educators reported that Texas students were taught "most or almost all the information needed to answer all of the questions on the test."¹⁰⁷ Further, 78% felt that the mathematics portion was relevant to teaching success and 88% reported that the reading, writing, and essay portions of the test were also relevant to teaching success.¹⁰⁸ The

102. See *id.* at 388.

103. See *Debra P.*, 730 F.2d at 1412.

104. *Rebell*, *supra* note 6 at 388.

105. *Lulac*, 793 F.2d 636 (5th Cir. 1986).

106. See *id.* at 639.

107. See *id.* at 640.

108. *Id.*

consultants reported that the "the PPST met the requirements of validation for relevance to the teacher education program and to performance as a classroom teacher." The only counter argument to test validation was based upon expert testimony that the test was biased against minority students. The court held that, based upon teachers' opinions, the PPST did determine whether applicants had basic skills essential both to success in passing pre-certification education courses and performing proficiently in teaching.¹⁰⁹ Using a combination of what appears to be instructional validity with face validity (does the test appear to measure what it purports to measure), the injunction against the use of the test was removed.

Even with the expense and difficulty of establishing instructional validity, the concept has wide application when the NTE is used for teacher certification. In its publication, "Guidelines for Proper Use of NTE Tests"¹¹⁰ ETS suggests that when the NTE program is used for credentialing, the certifying agency should:

Validate the tests to determine that they measure a representative sample of the knowledge and skills required for the appropriate certification and that any qualifying scores used differentiate between those who do and those who do not possess the requisite knowledge and skills.

Components for conducting a validation of NTE tests for certification and for providing potential qualifying scores should include the following:

a. An assessment of the appropriateness of the tests' content, given relevant teacher-training curricula, together with curricular analyses by appropriate experts, such as teacher educators and students at teacher-training programs in the state.

b. An assessment of the appropriateness of the tests' content, given relevant job requirements, with participation by practicing professionals, such as classroom teachers and school administrators.

c. An explicit process for establishing the minimum qualifying score for certification, with participation in the estimation of required minimal levels of knowledge by practicing professionals, such as teacher educators, classroom teachers, and school administrators.¹¹¹

Many states¹¹² are now including instructional validation for competency tests to avoid potential legal challenges.¹¹³ However, the whole notion of instructional validity, that is mirroring teacher education studies to what the NTE measures, raises several controversies. At the least, it creates homogeneity within the teacher education community that has

109. *See id.* at 645.

110. NTE POLICY COUNCIL, GUIDELINES FOR PROPER USE OF NTE TESTS (1985).

111. *Id.* at 3.

112. *See, e.g.*, EDUCATIONAL TESTING SERVICE, REPORT ON A STUDY OF THE NTE CORE BATTERY TESTS BY THE STATE OF NEW YORK (1983 & 1985); EDUCATIONAL TESTING SERVICE, CONTENT VALIDITY ANALYSIS OF THE NTE SPECIALTY AREA TEST IN SPECIAL EDUCATION (1985).

113. Rebell, *supra* note 6 at 391.

uniformly patterned its curriculum to insure that students will pass competency tests. Proving instructional validity also negates the larger issue of whether standardized tests can be used to predict which candidates are more likely than others to become effective teachers. Given the enormous range of teaching contexts and the wide variety of instructional areas, how likely is it that a uniform, standardized test will have validity for the diversity of teaching assignments? Perhaps it is more honest to say that use of the NTE will screen for the prerequisite literacy skills and theoretical knowledge base that are necessary for functioning in a complex and demanding profession, rather than to attempt to validate its use through large-scale teacher opinion polls.

E. Challenges Using an Item Bias Approach

The widespread use of standardized testing for teacher certification has revealed trends that potentially point to bias against minority group members in the tests themselves and has raised concerns about the fairness of the instruments in use.

In most states, there are variations in aggregate performances of different groups in the population. In particular, the performances of black, Hispanic, and white examinees typically differ from one another. There is, of course, considerable overlap in the distributions of scores across groups, but the differences are significant enough to have raised concern in both the education community and the testing industry. The challenge to the testing industry and program administrators is to ensure that such differences are not the result of bias in the test instruments themselves.¹¹⁴

Recent commentators have criticized the tests in widest use as being "biased, ambiguous and ineffective in measuring critical teaching skills."¹¹⁵

An unbiased test "provides fairness in selection to a group of applicants who, if selected, would be able to succeed regardless of group membership."¹¹⁶ To remove test bias the technique of item bias analysis has recently been proposed by those who question the fairness of standardized tests for minority group members. It is their contention that employment tests should be developed so that there will not be disproportionately high failure rates for minority test takers. Item bias analysis results in the elimination of test questions where there is shown to be significant differences in pass rates based upon group membership. Critics of the techni-

114. ALLAN, NASSIF & ELLIOT, *BIAS ISSUES IN TEACHER CERTIFICATION TESTING* ix (1988).

115. *Teacher Tests Assailed as Biased and Vague*, N.Y. Times, March 16, 1988, at B7, col. 1.

116. Johnson, *Validity and Bias in Teacher Certification Testing*, in *Testing for Teacher Certification* 35 (1988).

que fear that its use has the potential for undermining the maintenance of competency standards in teacher certification tests, saying that "[t]his new concept could have serious adverse effects on the integrity of testing."¹¹⁷

"Virtually all standardized certification tests, whatever the form of their validation, have tended to have a substantial adverse impact on minority candidates."¹¹⁸ For example, fifty percent of the graduates of Virginia's five traditionally black colleges, applying for teacher certification since 1982, would have failed at least one of the three NTE Core Battery tests. The failure rate for graduates from white institutions in Virginia was 3.1%.¹¹⁹ In a 1983 North Carolina study, 41.3% of black students failed at least one of the NTE Core Battery tests, as compared to 1.7% of white students. Implementing the revised standards for North Carolina in July 1986, resulted in 56.8% black failures as compared to 4.2% for white examinees.¹²⁰ A recent study that analyzed data from 19 states, including New York, Connecticut, New Jersey and California, also revealed extreme differences in passing rates when the results were analyzed by group membership. The average passing rate for black examinees was from 15% to 50%, Hispanic examinees from 39% to 65%, Asian-Americans examinees from 37% to 77%, and white examinees from 71% to 96%.¹²¹ A table of results for the October 1986 administration of the Core Battery portion of the NTE in New York is provided in Table 1.¹²² Examination of these results substantiates disparate impact on minority examinees in a northern state without a history of a dual school system. No national data is available that depicts the overall extent of this phenomenon, but significant disparate impact that results in a small pool of minority teachers has been frequently commented on.¹²³ Data regarding specific disparate impact within states contracting with ETS can be obtained from the state teacher certification agency in a state summary report. Clearly, group membership appears to have a significant impact upon a candidate's probability of becoming a certified teacher if standardized tests are used as part of the certification process, ultimately resulting in a smaller pool of new minority teacher candidates.

The item bias approach has been used to fashion a remedy to remove test bias. The technique was used by the Court in *Golden Rule Insurance*

117. Rebell, *supra* note 6 at 391.

118. *See id.* at 392.

119. Jaeger, *Policy Issues in Standard Setting for Professional Licensing Tests in Testing for Teacher Certification* 192-93 (1986).

120. *Id.*

121. N.Y. Times, *supra* note 115.

122. Table provided by New York State Education Department, Bureau of Teacher Certification, Albany, N.Y. 12234.

123. Rubenstein, *supra* note 13, at 31.

**Table 1. NUMBER TAKING AND PASSING AND PERCENT PASSING OCTOBER 1986
NTE CORE BATTERY TESTS IN NEW YORK**

POPULATIONS TESTED	COMMUNICATION SKILLS (650)			GENERAL KNOWLEDGE (649)			PROFESSIONAL KNOWLEDGE (646)		
	# Tests	# Pass	% Pass	# Tests	# Pass	% Pass	# Tests	# Pass	% Pass
Examinees in State	4,740	3,792	80	5,147	3,751	73	4,031	3,441	85
Black	273	127	47	316	126	40	209	124	59
Mexican American	4	1	25	5	4	80	2	1	50
Native American	20	15	75	20	17	85	20	15	75
Asian	33	14	42	27	16	59	27	13	48
Puerto Rican	97	38	39	111	33	30	85	48	56
Other Hispanic	50	25	50	57	21	37	37	23	62
White	3,834	3,283	86	4,095	3,230	79	3,251	2,930	90
Other	76	51	67	77	56	73	63	46	73
No information	353	238	67	439	248	56	337	241	72
Male	954	771	81	925	790	85	790	644	82
Female	3,761	3,008	80	4,160	2,920	70	3,170	2,740	86
No information	25	13	52	62	41	66	71	57	80

*Co. v. Washburn*¹²⁴ as well as referred to by the court in *Allen v. Alabama State Board of Education*¹²⁵ to eliminate test questions from exams that resulted in disparate, minority pass rates. In *Golden Rule* five people who failed a test for insurance licensing brought suit against the Illinois Dept. of Insurance and ETS who had created the test. Citing a disparate pass rate for black test-takers ETS agreed to analyze test items by the percentage of discrimination between black and white examinees. The comparison pass rate for the life insurance test was 78% for white examinees and 65% for blacks, and 82% for whites as compared to 55% for blacks on the accident and health insurance test.¹²⁶ ETS agreed to modify the test and in its revision only used items on which blacks as a group performed as well as whites before using items where there was greater rate of discrimination. Caution as to the resulting test validity seems indicated. "If items are discarded on the basis of differences in item difficulties between blacks and whites, it is possible that those items not used will eliminate some of the knowledge domain the test is designed to measure."¹²⁷

ETS has attempted to minimize the significance of their settlement in *Golden Rule*, but plaintiffs have "heralded it as a major breakthrough," predicting wide application to such other commonly used tests as the Scholastic Aptitude Test (for undergraduate college admission) and the Graduate Record Exam (for graduate school admission).¹²⁸

This approach has already had a more liberal application in testing for teacher certification. In *Allen v. Alabama State Board of Education*¹²⁹ a new certification process was developed in the context of a challenge to the certification standards previously used. A new subject area examination for teachers was to be developed. One significant provision was the use of item bias analysis and by agreement, items which resulted in racial differences greater than 15% would not be used.¹³⁰

The procedure results in higher pass rates for minority candidates, but it does not further the cause of ensuring that those who pass competency tests are knowledgeable and well prepared for their profession. Test content may become distorted when item bias analysis results in the preference

124. *Rebell*, *supra* note 6, at 392 n.54 (citing *Golden Rule Ins. Co. v. Washburn*, No. 419-76, (Ill. Cir. Ct. 7th Jud. Cir. Nov. 20, 1984) (Consent Decree)).

125. 612 F. Supp. 1046 (M.D. Ala. 1985), *vacated*, 636 F. Supp. 64 (M.D. Ala. 1986), *rev'd*, 816 F.2d 575 (11th Cir. 1987).

126. *Rebell*, *supra* note 6 at 393.

127. Berk, *Judgmental and Statistical Item Analysis of Teacher Certification Tests* in *Testing for Teacher Certification* 173 (1986).

128. *Rebell*, *supra* note 6 at 395.

129. *Allen*, 636 F. Supp. at 65.

130. *Id.*

for including easy items. If all groups score correctly on an item, it does not have item bias. "Accordingly difficult concepts will be measured less often, even if such concepts are an important measure of ability to perform a job."¹³¹ It is also predicted that use of the item bias approach will create antagonism by white candidates through the elimination of items on which they do well. This is especially true in light of the finding that the eliminated items appear to be lacking "any apparent culturally or racially biased wording or content."¹³² Several prominent black leaders have commented against this trend because "maintenance of educational standards and of the integrity of certification tests will benefit both minority students and minority teachers."¹³³ Item bias approach has been used for affirmative action purposes, but its use has been widely criticized and caution has been expressed that it should not be applied to "justify legislative proposals that go far beyond the very limited term of the original agreement."¹³⁴ In the context of teacher certification testing where "the nature of the information being tested [is] relatively straightforward," where the passing criterion is set at the minimal competency level, and where the overriding goal is to insure that only qualified candidates will be certified, "the 'cultural unfairness' issues are of even less significance."¹³⁵ One commentator has predicted that in spite of the public policy concern for affirmative action programs "it appears that the courts ultimately will not insist on affirmative action requirements that are shown to be 'eroding legitimate standards.'" ¹³⁶

F. Example of Complex, Current Litigation

An examination of current litigation in Alabama provides insight into the complexity of issues raised by current challenges to teacher certification testing. In 1972 the Alabama State Board of Education adopted a resolution that developed standards to improve teacher competency.¹³⁷ By 1979 standards for certification were revised so that candidates for initial certification would be evaluated on the basis of three criteria: 1) course work and grade-point average; 2) a full-time field internship in the area of teaching specialization; and 3) standardized, externally prepared, administered and scored examinations in the field of specialization. In

131. Rebell, *supra* note 6 at 394.

132. *Id.*

133. Rebell, *supra* note 6 at 398.

134. Rebell, *supra* note 28, at 11.

135. *See id.* at 18.

136. *See id.* at 12.

137. Alabama Department of Education, Alabama Initial Teacher Certification Testing Program: Registration Bulletin 1986-1987 2 (1986). [hereinafter Alabama Bulletin]

1981 certification testing was expanded to include testing candidates who were currently teaching but wanted to qualify for advanced level certification.¹³⁸

A set of criterion-referenced tests was custom designed for the state of Alabama by National Evaluation Systems (NES).¹³⁹

A criterion-referenced test is designed to measure a candidate's knowledge and skills in relation to an established standard of competence (criterion) rather than in relation to the performance of other candidates. The explicit purposes of these tests is to help identify those candidates for certification who have attained minimum standards of knowledge and competence required to perform satisfactorily in their fields of specialization.¹⁴⁰

The certification tests developed included a Basic Professional Studies Test (BPS) and 45 different tests for various areas of teaching specialization.

The tests were validated by several different processes: review by Alabama educators who were specialists in their content fields for relevance; development in accord with state-adopted curriculum; and by job analysis survey using 2,700 randomly selected Alabama teachers.¹⁴¹ According to the Director of Licensing and Certification at NES, custom designed tests have the advantage, among other things, of resulting in a significantly less disparate impact for minority candidates than does the National Teacher Examination.¹⁴² However, a class action suit was brought by four minority teachers and a predominantly black state university on behalf of all black people who have been or will be denied teacher certification by failure to pass the tests. Essentially, they are

138. Alabama's certification system involves the granting of Class B certification to those completing an undergraduate teacher education program and Class A or AA certification for those seeking advanced certificates. Other states such as New York call initial teacher certification "provisional" and the advanced certificate "permanent." Granting the advanced certificate usually entails graduate coursework (in New York the completion of at least a Master's Degree in a teaching or related field) and demonstrated competence as a teacher (in New York two years of satisfactory teaching experience).

139. National Evaluation Systems ("NES") is located in Amherst, Mass. and Palo Alto, Calif. They are the most experienced, national organization in the area of development of criterion-referenced tests for teacher certification, having developed more than 250 tests in more than 90 content area. The organization provides a complete range of testing services to state education agencies including current contracts in Alabama, Oklahoma, West Virginia, Connecticut and Texas. The range of services offered by NES includes: test design, validation, administration, reporting and evaluation. They also sponsor conferences on the topic of certification testing. National Evaluation Systems, *Improving Education Through Certification Testing 1 & 12*.

140. Alabama Bulletin, *supra* note 137 at 4.

141. *Id.*

142. Telephone interview with Scott Elliot, Ed.D., Division Director, Certification and Licensing NES (May 18, 1987).

challenging these custom designed tests for impermissibly discriminating against blacks.

Initially, the parties to the lawsuit were found to have agreed to a settlement that was approved by the court. However, on November 4, 1985, the Alabama State Board of Education moved for a rehearing, seeking to have the court reconsider its original finding that the parties entered into a binding settlement.¹⁴³ The nature of the settlement would have created a total restructuring of Alabama's certification process and it would have also entailed a substantial expenditure of state funds over and beyond the funds expended to develop and validate the new testing program. The proposed settlement was highly controversial and expensive.

The settlement created several novel aspects to the certification program. In addition to use of scores on the newly developed NES examinations, the candidates' grade point averages would be used with the test scores. If a candidate did not make the required cut-score, the grade point average would be counted 50% and test scores would be counted 50% to determine certification eligibility. Additionally new subject area tests would be developed with several specific provisions: (1) they would meet applicable professional guidelines; (2) development and administration would be reviewed and monitored by a three person expert panel appointed by the court after nomination by the parties; and (3) test items would be subject to limits of permissible discrimination by race. Item bias analysis was incorporated so any item that resulted in a racial difference greater than 15% would not be used. Moreover, if the above provisions did not result in a black pass rate equal to 90% of the white pass rate, an added number of the most highly qualified black candidates would be certified to make the differential pass rate come within 10% by racial classification.¹⁴⁴

The course of the Alabama litigation is quite complex and resulted in Alabama's certification process being contested for six years. In the first lawsuit, July 3, 1985, the United States District Court found that the settlement was agreed to by the parties but was not signed. Nevertheless, the court held that the agreement should be enforced.¹⁴⁵ A motion for rehearing was subsequently filed as to whether the agreement, reduced to a writing but unsigned, could be enforced by the court. In February of 1986 the same United States District Court found that the agreement could not be enforced.¹⁴⁶ Considerable weight was given to the fact that the Board of Education was unwilling, at that time, to institute the changes specified

143. *Allen*, 636 F. Supp., at 65.

144. *Id.*

145. *See id.* at 1048.

146. *Allen*, 636 F. Supp. at 65.

in the agreement. There was mounting public pressure against the agreement led by a broad attack by the State Attorney General.¹⁴⁷ Additionally, enforcing the agreement would result in a broad and substantial impact on the public, requiring a total restructuring of the state certification process at great expense.¹⁴⁸ Ultimately, the federal court was unwilling to intrude in a state's operations and enforce a settlement merely on an oral representation.¹⁴⁹

In April of 1987 the United States Court of Appeals reversed the previous court and reinstated the original consent decree. The Court held "that, in civil actions invoking Title VII, state defendants lack eleventh amendment protection and that otherwise valid consent-decree settlements agreed to in such cases by state agencies are enforceable notwithstanding the existence of the eleventh amendment."¹⁵⁰ Alabama's State Board of Education had to decide whether to appeal this decision to the United States Supreme Court. At one point the state was unable to describe the process by which it would certify teacher candidates.¹⁵¹ A final decision was reached in February 1988 when a United States District Judge approved of a plan proposed by the state that would award \$500,000 in damages among a potential 1,200 black candidates who were not certified because they failed the certification exam.¹⁵² A lawyer for the Alabama State Education Department reported that 900 blacks who had failed the original certification test passed the examination after their tests were rescored under the terms of the agreement. Each would receive approximately \$400 of the damages. Approximately 320 minority individuals who failed the test are still uncertified because they were deficient in meeting other certification requirements but would also receive damages if they were able to meet the requirements by August 30, 1988. After September 30, 1988 any additional money remaining would be evenly distributed among all who qualified for certification.¹⁵³

This case raises many perplexing problems regarding certifying teachers by competency testing. The case was not decided on the merits of the proposed certification program, but was decided on procedural grounds. However, Alabama had made a good faith attempt to develop customized job-related tests, independently administered and evaluated by an outside

147. *Allen*, No. 86-7215, 816 F.2d 575 (11th Cir. 1987).

148. *Allen*, 636 F. Supp. at 73.

149. *Allen*, 816 F.2d 575 (11th Cir. 1987).

150. *Id.* at 9.

151. Telephone interview with Barbara Fennel, Teacher Testing Advisor for the Alabama State Department of Education (May 18, 1987).

152. *Georgia, N.E.A. Settle Suit on Teacher Testing*, *Educ. Week*, March 9, 1988, at 17, col. 1 [hereinafter *Educ. Week*, March 9, 1988]

153. *Id.*

agency. Even this process resulted in extensive disparate impact against black candidates. A proposed settlement, that incorporated extensive use of the item bias approach, created mounting public pressure and was seen as compromising the basic purpose of teacher certification testing. While Alabama's State Board of Education is sensitive to the issue of disparate impact on minority teachers, it had to weigh this concern against certifying unqualified candidates and thereby lowering the quality of education. A well recognized litigator in the field has commented, "the public interest and ultimately the interest of the ill-prepared minority candidates themselves would be better served by pursuing reforms that would maintain the integrity of the tests."¹⁵⁴

Perhaps because of the extensive litigation in Alabama, the state of Georgia in two years agreed to settle a federal lawsuit brought by the Georgia affiliate of the National Education Association (NEA) involving the Georgia Teacher Certification Testing (TCT) Program.¹⁵⁵ The Georgia plaintiffs argued that the TCT was both racially biased and was not a valid predictor of performance as a teacher. The remedy in this case is rather unique for it allows Georgia to continue using the TCT as a screening device for granting initial certification and for the renewal of certification for practicing teachers. Thereby, the Georgia Association of Educators has acknowledged the validity of the TCT and has agreed, along with the NEA, not to challenge the validity of the test in the future. The state, too, has granted concessions. It will revise all subject-area certification tests by 1991, will provide for the development and delivery of TCT study courses, will provide study grants ranging from \$2,000 to \$6,000 per individual, and will administer three free testings to qualified applicants. The state school superintendent feels that this plan will benefit everyone because the state needs more qualified teachers and that those individuals who use the study grants will become more knowledgeable teachers and are likely to be employed in the school system. The superintendent feels that this pragmatic solution results in "all [being] winners" because losing such a lawsuit would be a multi-million dollar loss to the state, and that resources are better allocated towards improving the ability of teaching candidates and improving the certification testing program.¹⁵⁶

That standardized teacher certification tests result in adverse disparate impact among minority test-takers is apparent. Some have blamed the

154. Rebell, *supra* note 6, at 397.

155. Educ. Week, March 9, 1988, *supra* note 152, at 1; Georgia Ass'n of Educators v. State of Ga., civil action #1:86-cv-2234-cam (N.D. Ga. 1986).

156. Georgia Dept. of Educ., State of GAE Settle Teacher Testing Case, press release (Sept. 1988).

tests themselves, and others have looked at the ability of the test-takers. The most persuasive solution towards remedying this crisis appears to be proposals that provide extensive remedial help to strengthen the candidates' skills. It has been suggested that upon failing an initial exam candidates should be informed about their individual strengths and weaknesses and that remedial assistance be provided to address the skill deficiencies.¹⁵⁷ Job analysis procedures should be applied so as to develop tests that are valid for their intended purpose.¹⁵⁸ All who are involved in the development of tests should be made aware and sensitive to bias issues in test development — language usage, stereotyping, representational fairness (with attention to occupation, activity, and gender among other factors), and content inclusiveness.¹⁵⁹ There is much to be gained from creating better tests, strengthening the skills of teacher candidates and attempting to ensure diversity in the teaching profession.

5. Alternative Models to Current Certification Practices

The issues raised by certification testing include disparate impact on minority members, problems with test validity, and litigation that challenges its use. There is no question that testing for certification will continue to be common practice in the effort to raise educational standards. Current testing practices result in a combination of one or more forms of evaluation in four areas: basic skills testing (utilized by 34 states); professional skills testing (utilized by 25 states); on-the-job analysis (13 states); and content knowledge testing in the specialty area (26 states).¹⁶⁰ States wishing to insure both the legality and quality of their testing programs have several options to choose from including use of off-the-shelf standardized testing (usually the NTE), or the creation of a customized testing program. This section will describe two alternate models to the use of the National Teacher Exam, customized test creation and the very recent process evaluation method discussed in a proposal by the research group at Stanford University.

A. *The Use of a Customized Testing Program.*

A state can contract for the development of a certification test program on a customized basis. Essentially, this process involves the development

157. Rebell, *supra* note 6, at 399.

158. Elliot, *Preventing Bias in Teacher Certification Tests Through Valid Job Analysis Procedures*, in *Bias Issues in Teacher Certification Testing* 69 (1988).

159. *Bias Issues in Test Development*, in *Bias Issues in Teacher Certification Testing* 117 (1988).

160. Pecheone, Tomala & Gorgione, *Building a Competency Test for Prospective Teachers in Testing for Teacher Certification* 99 (1986).

of a measurement device keyed to objectives determined by the state and validated through a job analysis survey within the state.¹⁶¹ The customized test approach is now being used in Texas, Georgia, Alabama, West Virginia, Illinois and Connecticut. Texas, the largest consumer of a customized test program, provides an excellent case study for the development and implementation of such a testing program.¹⁶²

In 1981 the Texas legislature passed a bill which required that anyone seeking teacher certification perform satisfactorily on comprehensive examinations. The State Board of Education chose to have a testing program developed to meet the state's specific certification requirements. The Texas Education Agency (TEA) contracted with NES to design and develop 63 separate tests that would constitute the Examination for the Certification of Educators in Texas (ExCET) Program.¹⁶³ Since the program's inception it has been subjected to both legal and technical evaluations and has won the respect of educators both in and out of the state.¹⁶⁴

The Texas program is one of the most extensive criterion-referenced certification programs because all tests were developed to evaluate the skills and knowledge considered to be essential for effective teaching. The program was created to evaluate entry level skills expected of beginning teachers. Tests were developed for each of 60 different teaching specializations and for three general professional areas. Thousands of Texas teachers and students provided information about their School experience to form a job-related foundation for the development of the program. Over one thousand Texas educators, including classroom teachers and faculty from schools of education, consulted, reviewed and revised the test items as part of the extensive validation study.

Review of Table 2 demonstrates the elaborate, multi-phase process of the test development process and validation.¹⁶⁵ The process extensively in-

161. Gorth, *supra* note 5, at 4-5.

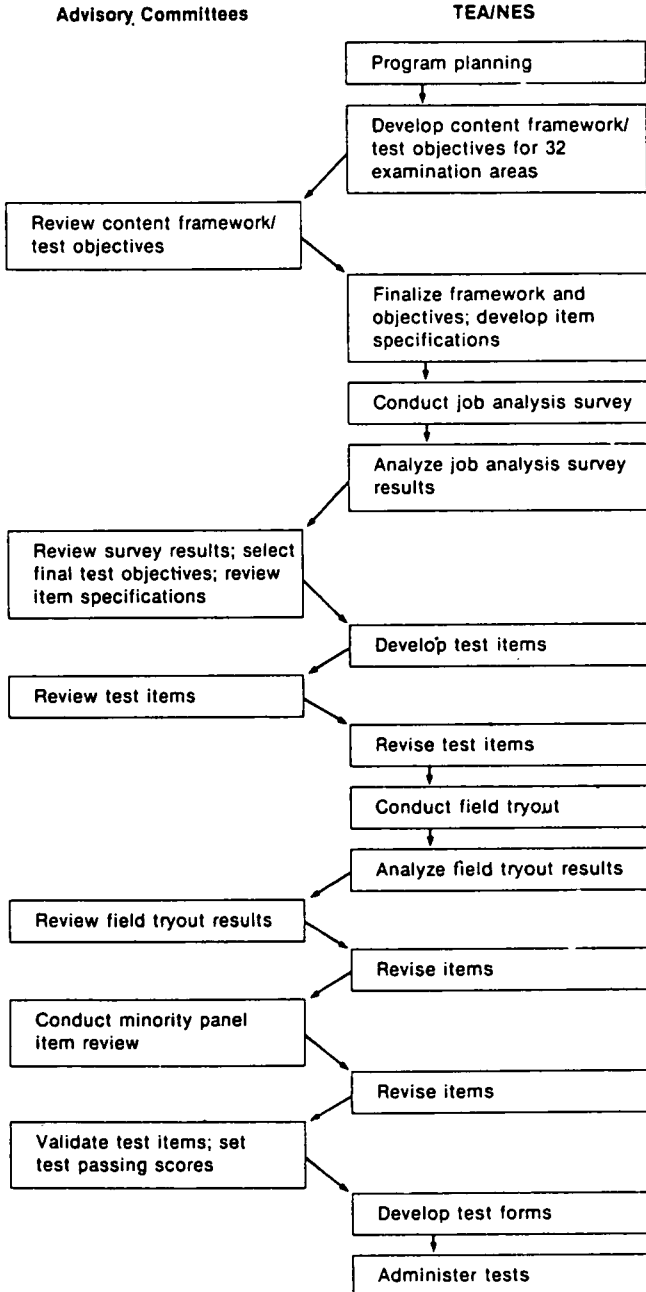
162. The author gratefully acknowledges the cooperation of the Texas Educ. Agency that supplied a number of published and draft documents describing the certification process in Texas. The information that follows is based on a draft document: NATIONAL EVALUATIONAL SYSTEMS, EXCELLENCE IN EDUCATION: EXAMINATION FOR THE CERTIFICATION OF EDUCATORS IN TEXAS (EXCET) PROGRAM (1987). This pamphlet was released by the office of Teacher Assessment, Texas Education Agency, Austin, Texas. [hereinafter Texas Bulletin]

163. Most recently Texas lawmakers have repealed a section of the 1984 reform that would have required all public school teachers to take tests measuring mastery of the subjects they teach. The tests would have been administered to approximately 210,000 presently teaching in Texas. It is reported that Texas feels that the current testing program is sufficiently extensive to ensure quality education and does not need to evaluate inservice teachers. *Subject Matter Tests for Teachers Repealed in Texas*, Educ. Week, May 6, 1987 at 11, col. 1.

164. Texas Bulletin, *supra* note 162 at 1.

165. Table contained in Texas Bulletin, *supra* note 162.

Table 2.
Examination for the Certification of Educator in Texas (ExCET)
Test Development Process



volved the cooperation of teachers and teacher educators at every phase as well as validation techniques dependent upon field tryout. This results in tests specifically designed to meet the instructional goals of the Texas school curriculum as well in tests that have the approval of the state's practitioners.

The test program is cooperatively administered by TEA and NES several times a year in a standardized test environment. After testing, examinees are given score reports that provide diagnostic information about their performance. These score reports are shared with TEA and the teacher training institutions to provide feedback for educational review. Examinees are permitted to retake the examinations as many times as they wish, using their score reports as a basis for remedial work when needed.

In February 1987, 8,973 candidates took 17,531 tests. Approximately 85% of the tests taken were passed with these results: 64.3% Hispanics passed, 51.7% Blacks passed, and 90.1% "others" passed.¹⁶⁶Data was also compiled regarding the pass rates by institution and program. This provided information for the teacher training institutions in the state to be used for curriculum review and the establishment of remedial programs where indicated. It is also expected that the institutional pass rates will be used as a factor in students' choice of training institution.

The positive features of such a customized program include: (1) test program reflects the unique needs and requirements of Texas with a carefully controlled match between state objectives and test items; (2) test program provides essential information to support certification decisions; (3) test development and validity components involve an extensive commitment by inservice educators, teacher candidates, and teacher educators; (4) both examinees and teacher training programs are supplied with information to assist in preparation of future teachers; (5) tests appear to be equitable to all examinees; (6) tests provide a uniform and positive experience for the examinees; and (7) state policy makers are provided with information to assist in program and policy modifications and review.

While off-the-shelf tests such as the NTE provide economy in both cost and effort, it is likely that customized testing programs will result in greater satisfaction and a more valid appraisal of certification candidates. The customized test approach, incorporating all segments of the teacher education community, also appears to result in significantly less disparate impact to minority group members. The customized method of certifica-

166. NATIONAL EVALUATION SYSTEMS, REPORT ON RESULTS OF THE EXAMINATION FOR THE CERTIFICATION OF EDUCATORS IN TEXAS (EXCET) PROGRAM THROUGH THE FEBRUARY 1987 ADMINISTRATION, 1 (1987).

tion testing, no doubt, is both costly and time-consuming. However, its benefits appear to outweigh its costs.

B. *The Stanford Project*

The Carnegie Corporation's recent document, *A Nation Prepared: Teachers for the 21st Century*, has resulted in the funding of a research project at Stanford University to develop sample test exercises that will result in a better system for evaluating teachers than tests in current use. Ultimately, the project will create a national teacher test so that teaching candidates can be certified by a peer review board on performance based tasks. A panel of distinguished public figures, teacher educators and teachers has recently been appointed to serve as both an advisory and review board for the project. The director of the project, however, has distinguished the work of this group from that of state licensing boards who are responsible for the initial certification of teachers. He feels that the work of this group is directed towards the professional certification of those who have already been licensed by their states and voluntarily seek to attain a higher credential. In his opinion, states will continue to license teachers "in their own, often flawed, ways."¹⁶⁷ The creation of a new teaching credential with high standards is expected to elevate teaching to a true profession. Some are skeptical according to Albert Shanker, President of the American Federation of Teachers.

But many people doubt that teaching can ever become a real profession. They think there's an innate difference between teaching and other fields. No two professions are alike, and teaching will develop in its own way. But if one looks back at what medicine was like early in this century (Low prestige, poor pay, rival philosophies of medical practice, low or no standards of entry and scandalously poor medical education), or law or business administration, we can see that teaching faces no greater hurdle in becoming a full profession that did other occupations.¹⁶⁸

While the project is geared towards the creation of a national teacher test, a primary task appears to be a determination of the kinds of assessment to include. The first phase of the project's development is to create teacher competency tests in two specific areas: teaching fractions to elementary school students and teaching the American Revolution to high school students. This narrow focus was chosen as a demonstration that a new form of assessment can be created.

The work of the team is to create a performance-based set of activities

167. Hechinger, *About Education*, N.Y. Times, March 2, 1988 at B7, col. 1. [hereinafter N.Y. Times, March 2, 1988]

168. Shanker, *Where We Stand*, Educ. Week, May 27, 1987 at 12, col. 1.

that will result in a test that appears to look more like teaching than tests in current use. A proposed model has three separate tiers of evaluation: part one focusing on understanding the content of the subject matter; part two examining the capacities needed to teach that content; and part three being the direct observation of teaching by carefully trained observers. An example of one kind of assessment would be to have a candidate critically analyze a proposed textbook, because this is a task that teachers often do. Another proposed activity is to have the candidate review a videotape of his/her teaching and then to critically analyze the lesson objectives, the activities, and the quality of teaching.

Unlike multiple-choice tests in current use, this new form of evaluation is guided by the premise that teaching is complex and that there is no single correct way to teach. It is also felt that the best measure of teacher competency reflects the analysis of performance-based, process skills that teachers are expected to use everyday. On the new tests teachers will be given the opportunity to make decisions, and then explain and justify their choices. Current examinations focus upon generic knowledge-based competencies that are supposed to be transferrable from one teaching context to another, and the paper and pencil, multiple-choice format does not attempt to measure the kinds of dynamic qualities that make teachers either competent or inadequate.¹⁶⁹ Additionally, the proposal attempts to elevate the measurement of teacher competency from what is currently perceived to be only a minimal competency assessment of basic literacy and general knowledge to the assessment of competencies that are more truly reflective of what makes for superior teaching.

Recognizing the disparate impact of current evaluation systems on minority candidates, the Stanford team has speculated that "one reason so many minority candidates fail teacher tests is that the exams are far removed from classroom situations."¹⁷⁰ How minority groups will perform on certification evaluations appears to be an essential concern of this approach, because "if this whole standards setting and assessment process further exacerbates the already unacceptable circumstances of minority participation in the teaching force, then the results will be totally unacceptable."¹⁷¹

It is predicted that it will take at least 4 years before a teacher candidate will be able to undertake this new form of certification. Multi-level teacher evaluation systems probably will be used to differentiate between practitioners at various levels of experience and competency. Currently, about a

169. N.Y. Times, March 2, 1988, *supra* note 168.

170. *Stanford Project*, *supra* note 15 at 18, col. 3.

171. *Id.*

dozen different types of teaching exercises are being developed and piloted with teachers in Oregon, California, Connecticut, Michigan and Pennsylvania. The results of these studies are eagerly anticipated by a profession that has become very critical of evaluation systems now in place.

6. Summary and Conclusions

This article surveyed and analyzed legal challenges to certification based on the wide-spread use of competency tests. Recent trends including the concern over disparate impact on minority group members and the item bias analysis approach were also discussed. The article attempted to portray the complex situation faced by state education agencies that have been trying to elevate teaching competency through assessment of prospective teachers and at the same time attempting to insure diversity within the work force.

Customized testing programs and the current research at Stanford University were described and suggested as alternatives to the use of off-the-shelf tests. It is anticipated that the implementation of either of the alternative models will result in greater satisfaction by teaching candidates and state certification agencies thereby obviating much of the current litigation.

In spite of the negative effects caused by the litigation in this area, issues have been raised that are expected to contribute to more effective teaching. Among these positive trends noted are:

- concern over instructional validation for any form of teacher assessment to ensure fairness to prospective candidates.
- on-going review of teacher education curriculum
- feedback mechanisms created so that job validation studies result in modifications in teacher education programs
- implementation of remedial programs for minority candidates who have had greater difficulty meeting with success on certification evaluations
- adequate notice given to candidates in states implementing new forms of certification evaluations
- implementation of test preparation programs with particular emphasis on remedial assistance
- implementation of pre-professional evaluation systems to diagnose potential strengths and weaknesses of teacher candidates well before certification testing is conducted
- analysis of test item weighing systems (item bias approach) and the resulting concern over test validity
- involvement of various components within the education community, including teachers, students, and teacher educators, in the certification reform
- shift to criterion referenced as opposed to norm referenced evaluation systems that results in identification of critical teaching behavior rather than comparing individuals

ETS has announced that a new form of assessment is being created to replace the NTE. The new test is expected to be available in 1992 and will be radically different from the form of the test in current use. ETS has acknowledged that educators agree that a "new and better kind of test for licensing teachers was needed."¹⁷² The new test will evaluate teachers at three phases of their education and training: during the sophomore year in college, at the completion of college or a teacher education program, and after a supervised teaching experience. A variety of measures of evaluation will be used as part of the new test battery that will be more interactive and more closely related to actual teaching skills. ETS no longer expects to measure teaching competence on the basis of multiple-choice tests alone. They are hopeful that the new form of assessment may be of benefit to minority students who will have several different means by which to show their competence. There are practical considerations to be worked out before the test can be implemented and individual states will need to develop their own validation techniques and set cut-off scores. However, ETS' admission that it's time for a new form of teacher certification test is a significant event in the evolution of standardized testing for teaching certification.

Additionally, more attention is currently being focused upon insuring minority group representation in the teaching profession. At a time when the number of minority students is increasing in our nation's schools it is expected that the number of minority teachers is likely to fall sharply within the next decade. There are several reasons for this trend: a significant number of teachers express a desire to leave teaching within the next 5 years and this is particularly true for black and Hispanic teachers,¹⁷³ the number of black and Hispanic students choosing to enter the teaching profession appears to be decreasing as other higher paying professions actively recruit minority candidates, and if significant numbers of minority candidates choosing to become teachers are eliminated because of the disparate impact of standardized, paper and pencil, multiple-choice tests, there will be fewer minority teachers eligible for certification. Many viewing the shortage of minority teachers predict that the smaller numbers of minority teachers will drastically jeopardize attempts to upgrade urban education.

As the proportion of white teachers grows, role modeling that might encourage minority students to pursue careers in education decreases. And as the proportion of minority teachers falls, the perceived importance of academic achievement to

172. *Different Tests of Teaching Skill Planned by Firm*, Educ. Week, Nov. 2, 1988, at 1, col. 1.

173. *Many Minority Teachers Plan to Quit, Poll Finds*, N.Y. Times, October 5, 1988, at B13, col. 1.

minority students also declines. The presence of appropriate role models in schools is crucial for helping minority students see the academic environment as one in which they can succeed — and should strive to do so. Without close, living examples, many of these young people will continue to hold a sadly limited image of success — one that devalues academic achievement.¹⁷⁴

Programs that identify the reasons why minority candidates fail to qualify on teacher certification tests and provide needed remediation are to be encouraged. Although they cannot compensate for life-long deficiencies in the formal education of many minority candidates, they can create needed opportunities for minority candidates to qualify as teachers.

In summary, the current legal and regulatory environment has had a profound influence on the design and implementation of teacher certification programs and the teacher certification process. It is likely that the legal challenges have resulted in creating better competency tests, thereby insuring that those who teach are good teachers, and ultimately restoring public confidence in the profession of teaching. It has also heightened the concern for maintaining and insuring diversity within the teaching profession. This has created an environment in which remedies to the problem can be fashioned by those within the profession without resorting to the expense, time and adversarial relationships that litigation creates. The responsibility for solving this complex problem is more properly placed within the education community than within the legal community. Elevating standards for teaching and insuring that competencies are fairly and accurately measured will result in better education.

174. Loehr, *The 'Urgent Need' for Minority Teachers*, Educ. Week, October 5, 1988 at 32, col. 1.

