

TESTING FOR SPECIAL SKILLS IN EMPLOYMENT: A NEW APPROACH TO JUDICIAL REVIEW

The use of standardized testing procedures as criteria for making employment decisions¹ has recently been subjected to increasingly frequent attack in suits brought under Title VII of the Civil Rights Act of 1964² or the fourteenth amendment.³ Ironically, this attack has developed in spite of Title VII's approval of testing⁴ and its express rejection of racial quotas⁵ as a basis for hiring choices, and the Equal Employment Opportunity Commission's suggestion that testing is a valid means of implementing Title VII's policy of non-discrimination in hiring.⁶ These challenges have resulted in a stringent judicial review of testing

THE FOLLOWING CITATION WILL BE USED IN THIS NOTE:

AMERICAN PSYCHOLOGICAL ASS'N, STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTS (rev. ed. 1974) [hereinafter cited as STANDARDS].

1. The term "employment decision" includes the hiring of applicants for jobs, the choice between current employees for promotions, and inter-departmental transfers. Throughout this Note, the term hiring decision will be used as the model; the principles discussed are applicable to the other decisions as well.

2. 42 U.S.C. §§ 2000e *et seq.* (1970). Subsection 2(a) of the Act reads as follows:

Unlawful employment practices.

(a) Employer practices.

It shall be an unlawful employment practice for an employer—(1) to fail or refuse to hire or discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin; *Id.* § 2000e-2(a).

3. [N]or shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws. U.S. CONST. amend. XIV, § 1.

4. [N]or shall it be an unlawful employment practice for an employer to give and to act upon the results of any professionally developed ability test provided that such test, its administration or action upon the results is not designed, intended or used to discriminate because of race, color, religion, sex, or national origin. 42 U.S.C. § 2000e-2(h) (1970).

5. Nothing contained in this subchapter shall be interpreted to require any employer . . . to grant preferential treatment to any individual or to any group because of the race, color, religion, sex, or national origin of such individual or group on account of an imbalance which may exist with respect to the total number or percentage of persons of any race . . . employed by any employer . . . in comparison with the total number or percentage of persons of such race . . . in any community, State, section, or other area, or in the available work force in any community, State, section, or other area. *Id.* § 2000e-2(j).

6. 29 C.F.R. § 1607.1(a) (1975). The regulation provides in part:

It is also recognized that professionally developed tests . . . may significantly aid in the development and maintenance of an efficient work force and, indeed, aid in the utilization and conservation of human resources generally.

procedures which has halted, and probably reversed, the trend toward more widespread use of employment tests.⁷

This development could lead to one of three results, any of which would have enormous impact on employment practices in the United States. First, and most likely in the short run, is a reversion to the use of subjective evaluation in making personnel decisions. Because subjective evaluation is not as amenable to reasoned analysis as objective testing, its use creates increased opportunity for consideration of such prohibited factors as race or sex in employee selection.⁸ The courts, accordingly, are not likely to relax their scrutiny of subjective techniques, and employers will have to overcome difficult burdens to justify any process which has a disproportionate racial or sexual impact.⁹ Such a burden may, in turn, force employers toward a second result—the utilization of quota hiring systems in an effort to avoid litigation.¹⁰ A third possible consequence of stringent judicial review of employment testing, and clearly the outcome which the reviewing courts intend, would be for employers to allocate more attention and resources to the design and implementation of tests which will measure important job skills without having a discriminatory impact on any minority group.

In light of the crucial role that court challenges to testing will play in the evolution of hiring practices, this Note will analyze the present judicial and administrative standards for assessing the validity of employment tests for jobs which require specialized training and suggest a new framework for determining when the various procedures available for validating tests must be applied.

7. Lancaster, *Failing System: Job Tests Are Dropped*, Wall Street J., Sept. 3, 1975, at 1, col. 6. In 1969, fifteen to twenty percent of all charges filed under Title VII involved testing. Cooper & Sobol, *Seniority and Testing Under Fair Employment Laws: A General Approach to Objective Criteria of Hiring and Promotion*, 82 HARV. L. REV. 1598, 1637 (1969). The use of testing had burgeoned in the period following World War II. See A. ANASTASI, *PSYCHOLOGICAL TESTING* 428-34 (3d ed. 1968); *Developments in the Law—Employment Discrimination and Title VII of the Civil Rights Act of 1964*, 84 HARV. L. REV. 1109, 1120 (1971).

8. For an example in the context of employment interviews, see *Shack v. Southworth*, 521 F.2d 51, 54-56 (6th Cir. 1975).

9. The EEOC definition of "test" is broad enough to encompass subjective techniques. Included are ". . . specific qualifying or disqualifying personal history or background requirements . . . , scored interviews, biographical information blanks, interviewers' rating scales, scored application forms, etc." 29 C.F.R. § 1607.2 (1975).

10. Note, *Business Necessity Under Title VII of the Civil Rights Act of 1964: A No-Alternative Approach*, 84 YALE L.J. 98, 106 (1974); see Fiss, *A Theory of Fair Employment Laws*, 38 U. CHI. L. REV. 235, 256, 279 (1971). Overt racial quota hiring systems have actually been adopted by some employers. See, e.g., *Hiatt v. City of Berkeley*, 10 F.E.P. Cas. 251 (Cal. Super. Ct. 1975).

TESTING FOR DIFFERENT PURPOSES:
THE NEED FOR A FLEXIBLE STANDARD

Establishing a *prima facie* case of discrimination under Title VII requires only that a minority group member show that test usage has resulted in a "racially disproportionate impact."¹¹ In many situations, this will not be difficult, since blacks and other minority group members have tended to score lower than whites on standardized written aptitude tests.¹² Once a racially disproportionate impact has been demonstrated, the burden shifts to the employer to justify his hiring practice.¹³ If he cannot show that the test is "job related"¹⁴ under the standards promulgated by the Equal Employment Opportunity Commission (the agency charged with administering Title VII), then he will not be permitted to use the test results in his hiring decisions.¹⁵

In situations not covered by Title VII,¹⁶ suits may be brought

11. The phrase "racially disproportionate impact" is preferred by the courts not only because it is less pejorative, but also because it is more accurate than describing a test as "racially discriminatory." See notes 95-99 *infra* and accompanying text.

12. See *Boston Chapter, NAACP, Inc. v. Beecher*, 504 F.2d 1017, 1021 (1st Cir. 1974); *Arrington v. Massachusetts Bay Transit Authority*, 306 F. Supp. 1355, 1358 (D. Mass. 1969); *Cooper & Sobol*, *supra* note 6, at 1638-41; *Fox & Lefkowitz, Differential Validity: Ethnic Group as a Moderator in Predicting Job Performance*, 27 PERSONNEL PSYCH. 209 (1974).

The reason for this deficiency in performance has not been related to any lack of innate intelligence but rather primarily to the socio-economic realities of a history of economic, cultural and education deprivation to which the black race has been subject. *Arrington v. Massachusetts Bay Transit Authority*, 306 F. Supp. at 1358.

13. *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 802 (1973); 2 A. LARSON, EMPLOYMENT DISCRIMINATION § 50.10 (1975).

14. *Griggs v. Duke Power Co.*, 401 U.S. 424, 431 (1971).

15. The Equal Employment Opportunity Commission (EEOC) defines as discriminatory

[t]he use of any test which adversely affects hiring, promotion, transfer or any other employment or membership opportunity of classes protected by title VII . . . unless: (a) the test has been validated and evidences a high degree of utility . . . and (b) the person giving or acting upon the results of the particular test can demonstrate that alternative suitable hiring, transfer or promotion procedures are unavailable for his use. 29 C.F.R. § 1607.3 (1975).

16. The scope of Title VII is uncertain. State certification procedures may be subject to the statute even though the state is not employing those examined. In *Sibley Memorial Hosp. v. Wilson*, 488 F.2d 1338 (D.C. Cir. 1973), the Court of Appeals for the District of Columbia held that Title VII applied despite the absence of a direct employer-employee relationship. *Sibley* involved a sex discrimination suit brought under Title VII by a male private-duty nurse who alleged that the hospital refused to refer him to female patients. Under the hospital's system, eligible nurses were referred to patients, who in turn became the actual employers. Support can be found throughout Title VII and the EEOC Testing Guidelines for the proposition that the intent of Congress was to deal with more than the conventional employer-employee relationship. Thus, an employer covered by Title VII who forecloses an individual's access to employment

under the fourteenth amendment.¹⁷ Here the standard used by the courts to review tests and their usage is relaxed; the test must bear only a rational relationship to the end which is sought.¹⁸ Since a test which is given to all prospective employees and graded uniformly will not be discriminatory on its face, the stringent review demanded in racial discrimination cases brought under the equal protection clause of the fourteenth amendment does not come into play.¹⁹

Neither the strict statutory standard of review nor the relaxed fourteenth amendment standard takes full cognizance of the premises generally accepted by experts in the field of testing. Although tests and their implementation are susceptible to abuse,²⁰ most experts contend that a test which is properly designed and administered is superior to subjective methods of selection.²¹ Valid objective tests can be especially useful in the context of jobs which require substantial specialized training, such as teaching, accounting, architecture, or the practice of law or medicine.²² In this area, the judicial decisions do not reflect a clear understanding of the different types of testing procedures and the purposes served by each. Consequently, the courts have tended to examine the tests being reviewed in the abstract, and have failed to give adequate attention to the ways in which the test results are being used.

opportunities with another employer violates the statute. *Puntolillo v. New Hampshire Racing Comm'n*, 375 F. Supp. 1089 (D.N.H. 1974) (Racing Commission in violation of statute for failing to certify a jockey on the basis of national origin thereby foreclosing his opportunities to be employed by race horse owners). See also *Hackett v. McGuire Bros.*, 445 F.2d 442 (3d Cir. 1971).

17. The fourteenth amendment, by its terms, covers only action by the states. Title VII, as originally enacted, exempted governmental bodies from its coverage. In 1972, this exemption was deleted so that governmental units as employers are covered. See note 103 *infra*.

18. See notes 90-102 *infra* and accompanying text.

19. See note 95 *infra* and accompanying text.

20. See note 99 *infra* and the cases discussed therein.

21. Testing is generally considered preferable to any alternative by both employers and minority group members. See Wilson, *A Second Look at Griggs v. Duke Power Company: Ruminations on Job Testing, Discrimination, and the Role of the Federal Courts*, 58 VA. L. REV. 844, 873 (1972). But see Cooper & Sobol, *supra* note 6, at 1676-77; Note, *Legal Implications of the Use of Standardized Ability Tests in Employment and Education*, 68 COLUM. L. REV. 691, 743-44 (1968). Some authorities continue to consider objective testing the best means for reducing bias. See, e.g., Grossman, *The Equal Opportunity Act of 1972, Its Implications for the State and Local Government Manager*, 2 PUB. PERSONNEL MANAGEMENT 370, 376 (1973); Comment, *Employment Testing: The Aftermath of Griggs v. Duke Power Company*, 72 COLUM. L. REV. 900 (1972).

22. Jobs of this type not only require some minimum level of knowledge or skill, but frequently involve the employer's representations to the consuming public that his employees possess the requisite proficiency.

For example, an employer who seeks to use an aptitude or general intelligence test to rank job applicants in the order of their prospective job performance can appropriately be required to demonstrate that the test results yield accurate predictions of ultimate performance.²³ Where, however, an employer does not use test results to rank applicants in the order of predicted performance,²⁴ but seeks only to ensure that applicants possess the specialized knowledge which a particular job requires, it is not always appropriate to demand a showing that a high test score is predictive of a correspondingly high performance level.

An illustration of the second type of situation may be helpful. An applicant for a position as a high school history teacher, it may be assumed,²⁵ should have some knowledge of that subject. It is therefore reasonable for a local school board which is screening applicants for the job to require that they demonstrate the requisite knowledge of history on a properly designed achievement test.²⁶ Those who score highest on such a test of knowledge, however, may not necessarily prove to be the most effective history teachers when their classroom performances are assessed.²⁷ Many other factors, some of which are not readily measurable, contribute to the effectiveness of a classroom instructor; "[t]here currently are no good measures of patience, warmth, humor, love of learning."²⁸ Still, the fact that an achievement test cannot predict which of the tested applicants will prove to be the best classroom history teachers should not deprive the local school board of a device

23. See, e.g., *Boston Chapter, NAACP, Inc. v. Beecher*, 504 F.2d 1017, 1021-22 & n.9 (1st Cir. 1974); *Kirkland v. New York State Dep't of Correctional Serv.*, 374 F. Supp. 1361 (S.D.N.Y. 1974), *aff'd in part*, 520 F.2d 420 (2d Cir. 1975).

24. An employer's decision not to rank applicants according to test results could be prompted by several factors: (1) The test, in the employer's view, is not an adequate predictor of which employees will perform better on the job, although it is useful for other purposes, see note 37 *infra* and accompanying text; (2) Refining a test to make it predictive could be very expensive; (3) The employer may prefer other means for ranking those applicants who pass the test, such as extensive interviews or references.

25. Although it is assumed here that a knowledge of history is required in order to teach the subject, a properly conducted job analysis would be necessary to establish exactly what skills and abilities are required for any given position.

26. Achievement tests used in this manner are analogous to the requirement of a degree or diploma as evidence of proficiency in an area of specialized training. See note 84 *infra* and accompanying text.

27. It may be difficult to determine which teachers are "successful," even apart from the question of testing for employment decisions. Cf. Wilson, Dienst & Watson, *Characteristics of Effective College Teachers as Perceived by Their Colleagues*, 10 J. EDUC. MEASUREMENT 31 (1973); Subkoviak & Levin, *Determining the Characteristics of the Ideal Professor: An Alternative Approach*, 11 J. EDUC. MEASUREMENT 269 (1974). See notes 67-68 *infra* and accompanying text.

28. Huff, *Credentialing by Tests or by Degrees: Title VII of the Civil Rights Act and Griggs v. Duke Power Company*, 44 HARV. EDUC. REV. 246, 265 (1974).

for ensuring that its history teachers have an adequate knowledge of history, even where the test has a disproportionate impact on a minority group.²⁹

Thus, the purpose for which an employer uses a test is clearly relevant to the determination of whether a test which has a racially disproportionate impact should be permitted. To understand how this problem should be treated, it is first necessary to have a basic understanding of the principles of testing and to examine the statutory and administrative standards applied to employment testing devices.

Testing Theory: Criterion-Related and Content Validation

The concept of "validity" refers to the extent to which a test measures the qualities it purports to measure;³⁰ it is therefore a reflection of the appropriateness of inferences drawn from the results of a test.³¹ Two methods are commonly used to determine the validity of tests in employment situations.³² The first, criterion-related validation, can be used to assess the degree to which a test given by an employer serves the purpose of estimating how well an individual will perform in a particu-

29. In a slightly different context, the Fifth Circuit recognized the state's interest in establishing that applicants admitted to the Georgia Bar possess a minimum knowledge of the subjects taught in law school, *Tyler v. Vickery*, 517 F.2d 1089 (5th Cir. 1975), cert. denied, 44 U.S.L.W. 3719 (U.S. June 14, 1976), even though it was not asserted that those who achieve the highest passing scores on the bar examination would prove to be the best practicing attorneys; cf. *United States v. North Carolina*, 400 F. Supp. 343, 348-49 & n.6 (E.D.N.C. 1975) (three-judge court) (acknowledging state's right to adopt minimum knowledge requirements for teachers but requiring that state validate the cut-off score used to determine minimal competence).

30. SOCIAL SCIENCE RESEARCH COUNCIL, *SOCIAL EXPERIMENTATION: A METHOD OF PLANNING AND EVALUATING SOCIAL INTERVENTION* 32 (H. Riecken & R. Boruch eds. 1974).

31. STANDARDS 25. For a full discussion of the area of test validation, see A. ANASTASI, *supra* note 7; L. CRONBACH, *ESSENTIALS OF PSYCHOLOGICAL TESTING* (2d ed. 1960); M. DUNNETTE, *PERSONNEL SELECTION AND PLACEMENT* (1966); E. GHISELLI, *THE VALIDITY OF OCCUPATIONAL APTITUDE TESTS* (1966); R. GUION, *PERSONNEL TESTING* (1965); R. THORNDIKE, *PERSONNEL SELECTION: TEST AND MEASUREMENT TECHNIQUES* (1949); Note, *supra* note 21, at 696-706 (1968).

32. A third process, construct validation, is based on a different approach to the concept of validity. The purpose of this process is to determine what "constructs," or personal characteristics, distinguish those who score high on a test from those who score low. After extensive inductive study, the analyst formulates hypotheses about the characteristics which are being tested for. Independent testing is then conducted to determine whether those who score well in fact possess the suspected characteristics. See STANDARDS 29-30.

A fourth (and highly controversial) concept, "differential validity," is based on the idea that a test may be a valid predictor for one group but not for another. See Sandman & Urban, *Employment Testing and the Law*, 27 LABOR L.J. 38, 53 (1976).

lar job.³³ There are two categories of criterion-related validation: concurrent and predictive. A concurrent validation study compares test scores with evaluations of job performances of the tested employees.³⁴ Insofar as the test scores are found to be significantly correlated with successful job performance, the employer can then give the same test to future job applicants with a reasonable expectation of selecting those who will perform best as employees. There are problems inherent in this type of study; for example, many employment situations involve some form of post-hiring training, so present employees may not be comparable to job applicants.³⁵ The second type of criterion-related validation, predictive validation, is the conceptually sounder and preferred method, despite its greater cost in terms of time and money. Here, all applicants are given the test in question and hired regardless of their scores.³⁶ Subsequent job performance is then compared with the test results to determine the extent to which the examination predicts future success. This method is more nearly identical to the manner in which the employer ultimately intends to use the test.

Content validation, the second method commonly used to determine the validity of employment tests, is not a substitute for criterion-related validation. Psychologists recognize content validation studies as

33. Criterion validity measures the extent to which a particular test produces results which are consonant with some well-accepted indicator of whatever is being measured. For example, if an individual wants to measure alcoholic content of the blood using a breathalyzer, a blood test would provide the standard against which the validity of the breathalyzer test might be measured. A simple example of criterion validity in employment testing might involve the relationship of aptitude test score results to the actual output of a factory worker. Clearly, development of a standard is more difficult with employment tests where it is hard to decide what constitutes success.

34. Because concurrent validation employs the individual's present successful job performance as the chosen criterion, it can only serve as a measure of the status quo. STANDARDS 26.

35. Another example would be a situation where few of the present employees are minority group members. The test could be written in terminology or laden with values foreign to minority applicants and unrelated to future job performance; these spurious features would obscure the meaning of the test scores. As a result, many commonly used tests, especially those testing general aptitudes or culturally related achievements, have little if any value as predictors of performance in unskilled or semi-skilled jobs. *Developments in the Law, supra* note 7, at 1120-22. See also *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975); *Griggs v. Duke Power Co.*, 401 U.S. 425 (1971).

36. The hiring of *all* applicants is a critical difference. If only some of the applicants are hired, it is impossible to know what the test scores actually mean. If those employed were selected on the basis of the test, it is impossible to know whether those scoring lower on the test would actually have performed more poorly on the job. Alternatively, if those hired were selected on the basis of some criterion other than the test, it may remain unclear whether it is the test or that criterion which controls job performance. Only when all applicants are tested and examined can the validity of the test be properly evaluated.

appropriate only for achievement tests designed to measure an individual's present knowledge or skills.³⁷ Initially, the job must be analyzed and the area of specific knowledge required for successful performance must be delineated. A hiring test will be content-valid if it adequately measures only that area of knowledge, and no other. In the case of the history teacher discussed above, for example, the qualifying test would be content-valid if it accurately measured knowledge of the historical subject matter covered by the curriculum.³⁸

THE TITLE VII STANDARD OF REVIEW

Administrative Interpretation

Title VII of the Civil Rights Act of 1964 prohibits discrimination in employment on the basis of race, color, religion, sex, or national origin.³⁹ In theory, hiring must be done on a "color-blind" basis⁴⁰ so that any minority group applicant who is sufficiently qualified, or readily qualifiable, to perform a job is given an opportunity equal to that of a similarly situated white applicant. Testing is expressly recognized as a lawful employment practice, provided that the "test, its administration or action upon the results is not designed, intended or used to discriminate"⁴¹ In fact, discussion in Congress during the debate on Title VII focused upon salvaging professionally developed tests that measure relevant abilities and skills.⁴²

37. See A. ANASTASI, *supra* note 7, at 102, 391, 428-34; Cooper & Sobol, *supra* note 7, at 1643-44. It may, however, be appropriate in special circumstances to assess the content validity of aptitude tests. STANDARDS 28.

It is important to note that although a test may be content-valid, it is not necessarily suitable for predicting performance, particularly in positions for which on-the-job training is feasible. A content-valid test indicates only that the applicant has some proficiency in the tested skill or subject. Only where the tested area coincides almost entirely with the totality of skills required in a position would the test predict performance. Cf. STANDARDS 29.

38. Sources for determining the proper subject-matter of the test might include textbooks, course syllabi, and consultations with subject-matter experts. See A. ANASTASI, *supra* note 7, at 101.

39. See note 2 *supra*.

40. This would reject hiring by any quota-related method which mandated color-awareness. See 42 U.S.C. § 2000e-2(j) (1970), *quoted* note 5 *supra*; cf. *Hiatt v. City of Berkeley*, 10 F.E.P. Cas. 251 (Cal. Super. Ct. 1975).

41. 42 U.S.C. § 2000e-2(h) (1970). For the full text of this subsection see note 4 *supra*.

42. At the time of the Congressional debates, attention was focused on *Myart v. Motora, Inc.*, a decision by a hearing examiner of the Illinois Fair Employment Practices Commission. The text of this decision is quoted in full at 110 CONG. REC. 5662-64 (1964). The case was considered at the time to invalidate all professionally developed tests.

Debate over what constitutes "relevant abilities" has occupied many writers in the

The Equal Employment Opportunity Commission (EEOC)⁴³ has drafted "Guidelines on Employee Selection Procedures"⁴⁴ pursuant to the authority granted to it by Congress.⁴⁵ Stating that the Guidelines are based on a "belief that properly validated and standardized employee selection procedures can significantly contribute to the implementation of non-discriminatory personnel policies,"⁴⁶ the EEOC has sought to eliminate the use of testing as a vehicle for covert discrimination. The Guidelines, therefore, were carefully drafted to be both stringent and broad in scope.⁴⁷ For the purposes of the Guidelines, "test" is defined as "any paper-and-pencil or performance measure," including all formal, scored, quantified or standardized techniques such as measures of general intelligence, mental or learning ability, and knowledge or proficiency.⁴⁸ The Guidelines also treat specific qualifying require-

field. In a response to the arguments against employment tests presented in Note, *supra* note 21, Professor Ruch voiced the opinion that even a general intelligence test should be considered valid if general intelligence is relevant to the level to which the applicant might be promoted. Ruch & Ash, *Comments on Psychological Testing*, 69 COLUM. L. REV. 608, 611 (1969). *But cf.* Cooper & Sobol, *supra* note 7, at 1643:

It is sometimes argued that standardized intelligence tests are inherently related to business needs on the ground that every employer is entitled to prefer more intelligent employees This notion misconceives the function of tests. Industrial employers need people who can do industrial jobs better; to the extent that requires a certain mental capacity, the employer can be said to need a more "intelligent" employee or one with certain kinds of comprehension. But a paper and pencil test asking general questions does not necessarily measure the relevant mental capacity. It measures the capacity to answer questions on the test. This may or may not be related to the capacity to perform well on particular jobs.

43. The EEOC is composed of five members appointed by the President with the advice and consent of the Senate, with the members serving staggered five year terms. 42 U.S.C. § 2000e-4(a) (1970).

44. 29 C.F.R. §§ 1607.1 *et seq.* (1975). For a discussion of the development of the Guidelines and of the EEOC's very broad interpretation of Title VII, see Blumrosen, *Strangers in Paradise: Griggs v. Duke Power Co. and the Concept of Employment Discrimination*, 71 MICH. L. REV. 59 (1972).

45. 42 U.S.C. § 2000e-12 (1970). The Guidelines, ostensibly interpretations of Title VII, were not promulgated as rules under the notice-and-comment provisions of the Administrative Procedure Act, 5 U.S.C. § 553(c) (1970). Agency compliance with the statute is not required for purely interpretive regulations, *id.* § 553(b)(3)(A); however, several Guidelines provisions detail specific validation requirements not mentioned in the statute. These provisions might have elicited critical response from some psychologists had the notice-and-comment procedure been followed. *See* Humphreys, *Statistical Definitions of Test Validity for Minority Groups*, 58 J. APPLIED PSYCHOLOGY 1 (1973) (criticizing failure of Guidelines to set out methods of validation and resulting usage of statistically inaccurate procedures); Grossman, *supra* note 21, at 378 (EEOC Guidelines appear to raise a presumption against tests).

46. 29 C.F.R. § 1607.1(a) (1975).

47. For a discussion of the promulgation of the EEOC Guidelines by the person who served as Chairman of the Commission at the time of their drafting, see Blumrosen, *supra* note 44, at 60-61.

48. 29 C.F.R. § 1607.2 (1975).

ments, such as educational history, and other selection techniques, such as unscored or casual interviews.⁴⁹

The EEOC Guidelines expressly accept the American Psychological Association's *Standards for Educational and Psychological Tests*,⁵⁰ and the *Standards* in turn specifically recognize content validation to be appropriate for assessing the validity of achievement tests.⁵¹ The Guidelines, however, place great stress on criterion-related validation techniques for all employment tests,⁵² permitting resort to content validation only where the employer can show that criterion-related validation is "not feasible."⁵³

Since any test used as a basis for personnel decisions is predictive in some sense, it is reasonable that the Guidelines should express a preference for a validation technique which assesses predictive efficacy. Where test scores are utilized to rank applicants along a scale of anticipated job success, the predictive nature of the test is clear. Where the test is used only to ensure requisite skill or knowledge in a given area, the test scores

49. 29 C.F.R. § 1607.13 (1975).

50. 29 C.F.R. § 1607.5 (1975). This provision states that

[f]or the purpose of satisfying the requirements of this part, empirical evidence in support of a test's validity must be based on studies employing generally accepted procedures for determining criterion-related validity, such as those described in "Standards for Educational and Psychological Tests and Manuals" published by the American Psychological Association.

51. STANDARDS 28.

52. 29 C.F.R. § 1607.5. *Minimum standards for validation.*

(a) . . . Evidence of content or construct validity, as defined in ["Standards for Educational and Psychological Tests and Manuals"], may also be appropriate where criterion-related validity is *not feasible*. However, evidence for content or construct validity should be accompanied by sufficient information from job analyses to demonstrate the relevance of the content (in the case of job knowledge or proficiency tests) or the construct (in the case of trait measures). Evidence of content validity alone may be acceptable for well-developed tests that consist of suitable samples of the essential knowledge, skills or behaviors composing the job in question. The types of knowledge, skills or behaviors contemplated here do not include those which can be acquired in a brief orientation to the job. (Emphasis added).

53. *Id.* See, e.g., *Davis v. Washington*, 515 F.2d 956, 965 n.64 (D.C. Cir. 1975), *rev'd*, 44 U.S.L.W. 4789 (U.S. June 7, 1976); *Douglas v. Hampton*, 512 F.2d 976 (D.C. Cir. 1975).

Another federal agency, the Office of Federal Contract Compliance (OFCC), although purportedly imposing the same requirements as the EEOC, 41 C.F.R. § 60-3.1(c) n.1 (1974), accepts content validation for tests given to determine whether an applicant possesses some skill or knowledge needed for a job. The OFCC "Guidelines" provide:

(b) Content validity: Content validity is an appropriate type of validity to demonstrate for skills tests (that is, tests to determine whether an individual already possesses needed skills or knowledge) and certain other selection procedures but it is not ordinarily an appropriate type of validity to fully support aptitude or intelligence tests (that is, tests to determine whether an individual can learn needed skills or knowledge). Also, content validity sometimes may be accepted as provisional compliance where criterion related is preferred but is not technically feasible. *Id.* § 60-3.6(b).

are nonetheless predictive in the sense that they project a lack of success for those applicants who fail to achieve the minimum score. In addition, the only recognized justifications for the use of tests which have a disproportionate racial impact are the safety and efficiency of the employer's business.⁵⁴ That is, an employer can use such a test only if he can demonstrate that his business is made safer or more efficient through the selection of the best-qualified employees. The greatest business efficiency is achieved when the selection procedure actually indicates with some degree of precision which individuals among a given group of applicants will perform best on the job. A test which can demonstrate predictive criterion-related validity should therefore be preferred.

Judicial Review Standards

Under recent decisions of the Supreme Court, the EEOC Guidelines, including the marked preference for criterion-related validation, are to be accorded "great deference" in Title VII actions.⁵⁵ Neither the

54. See Note, *supra* note 10, at 108.

55. *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 421 (1975), quoting *Griggs v. Duke Power Co.*, 401 U.S. 424, 434 (1971); see *Sandman & Urban*, *supra* note 32, at 43. *Griggs* was the first case before the Supreme Court in which tests were challenged under Title VII. In *Griggs*, on the day Title VII became effective, a company with a history of overt racial discrimination began requiring a high school diploma and satisfactory scores on two professionally developed aptitude tests for its factory positions. 401 U.S. at 427-28. It was demonstrated that these requirements limited the advancement of blacks who were already working for the company, as well as the initial employment of black applicants. *Id.* at 426-28. Chief Justice Burger, writing for a unanimous Court, found that the employer had not shown the tests to be job-related in the face of their racially disproportionate impact. *Id.* at 431. In reaching this conclusion, the Court held that the general EEOC Guidelines under Title VII were "inescapably" entitled to great deference; the Court felt that the Guidelines were supported both by the language of Title VII and by its legislative history. *Id.* at 434-36. Despite the Court's interpretation, there are indications that the statute was never intended to prohibit tests administered equally to blacks and whites. See *Wilson*, *supra* note 21, at 852-58; *Developments in the Law*, *supra* note 7, at 1123-26; Note, *supra* note 21, at 706-10.

The *Griggs* Court paid deference only to the general Guidelines and not to the specific provisions regarding test validation. These technical validation Guidelines have been recognized as controlling at the Court of Appeals level. See, e.g., *Davis v. Washington*, 512 F.2d 956 (D.C. Cir. 1975), *rev'd on other grounds*, 44 U.S.L.W. 4789 (U.S. June 7, 1976); *United States v. Georgia Power Co.*, 474 F.2d 906 (5th Cir. 1973). The Guidelines have also been substantially adopted by the Office of Federal Contract Compliance. 41 C.F.R. § 60-3.1(c) n.1 (1974).

In *Moody*, a case arising out of facts very similar to those in *Griggs*, the Supreme Court, while recognizing that the Guidelines were not rules, did measure a test challenged under Title VII against the detailed test validation Guidelines. The company had a psychologist perform a validation study on the eve of the trial. In invalidating his findings, the Court relied heavily on their nonconformance with the Guidelines. 422 U.S. at 430-36. Chief Justice Burger, this time concurring in part and dissenting in part,

Supreme Court nor the lower courts, however, have attempted to construe the "not feasible" language of the Guidelines; the courts have yet to treat in a systematic manner the problem of determining when a content-valid achievement test is permissible in the absence of a showing of predictive validity. The Guidelines might be read to bar the use of content validation in every situation where an employer is unable to demonstrate the technological impossibility of designing a test which would have predictive validity. Alternatively, the "not feasible" language could support an interpretation that content validation is appropriate where it is either impracticable or very costly for an employer to develop a test which can be validated by the criterion-related approach.

The legislative history of Title VII⁵⁶ and the policy which the Act embodies both counsel against rigid rejection of a validating technique recognized as appropriate by the American Psychological Association. It must be recalled that Title VII does not automatically prohibit an employer from using a test which has a disproportionate impact on a minority group if the test can be justified by the added safety and efficiency which it brings to the employer's business.⁵⁷ With this in mind, one can proceed to examine the settings in which an employer might seek to use a test which has content validity but which has not been shown to have predictive validity.

CONTENT VERSUS CRITERION VALIDATION: A PROPOSED ANALYSIS

Predictive Ranking

Scores on a criterion-validated test, as noted above,⁵⁸ will show a positive correlation with some objective measure of successful job performance. Content validation, on the other hand, indicates only that the test does in fact measure the same skill or knowledge that it purports to assess.⁵⁹ In judging the permissibility of a content-validated test

id. at 449, could not accept the Court's deference to the EEOC. By using the Guidelines in this manner, he observed, the Court treated them as rules although they had not been promulgated as such under the Administrative Procedure Act. The Chief Justice drew a distinction between the Guideline relied upon in *Griggs*, which was a general interpretation of Title VII, and those followed in *Moody*, which delineated specific methods for proving job-relatedness. *Id.* at 451-52.

56. See note 42 *supra* and accompanying text.

57. See note 54 *supra* and accompanying text.

58. See notes 34-35 *supra* and accompanying text.

59. This point is an important one. It might be said that a test is not content-valid, even though it is an accurate measure of the subject it purports to measure, if that subject is not important in the job. There is really no difference between that statement and the

under Title VII, the critical question is therefore whether use of the test would contribute to the employer's efficiency in selecting capable employees. In determining efficiency, it must be ascertained whether the area of knowledge or skill tested for is in fact necessary or useful in the job which the employer seeks to fill. A properly designed and executed job analysis is normally used to answer the latter question. However, even where it can be shown that the test is an accurate measure of the applicant's knowledge of certain subject matter and that the subject matter is related to the job sought, the employer cannot claim that the test results are predictive. A content-valid test may be used to determine minimum competence, but it cannot support the inference that those who achieve higher scores will be the better performers on the job.

Thus, whenever an employer makes use of a content-validated test to rank applicants in the order of their test scores, it is appropriate that the employer bear a heavy burden in demonstrating why a criterion-valid test should not be required. The use of any test reflects an implicit belief on the part of the user that the test actually predicts which applicants will be the "better" employees. A content-valid test, however, is by definition an inadequate basis for ranking applicants. The employer is therefore confronted with a seemingly insoluble logical dilemma: there can be no business efficiency justification for using the content-validated test unless it is asserted that the test will fulfill a predictive function for which it is inherently unsuited.

An equally heavy burden should be imposed on an employer who does not use test scores for overt ranking purposes, but who sets a cut-off score in order to eliminate a selected percentage of applicants.⁶⁰ The use of such a cut-off, adjusted from year to year as market conditions demand, is in reality a crude form of ranking applicants. For example, should the employer set a cut-off score which he estimates would eliminate thirty percent of the applicants for a position, a score based on a projected need for approximately seventy percent of the applicant population, the employer is effectively using the test scores to predict that those above the cut-off will be more successful on the job than those below it.⁶¹

one in the text to the effect that a test may be content-valid but must still be related to job skills. The point worth emphasizing is that the content-valid test is not a predictor of success; it is rather a measure of proficiency in certain areas.

60. See *League of United Latin American Citizens v. City of Santa Ana*, 12 F.E.P. Cas. 651, 659, 673 (C.D. Cal. 1976); cf. *Walston v. County School Bd.*, 492 F.2d 919, 924-25 (4th Cir. 1974).

61. The fact that the cut-off score for one time period may be lower than that for the next belies any assertion that the cut-off is selected as the point which separates qualified from unqualified applicants.

Testing for Minimum Competence

Employment testing for jobs which require substantial specialized training presents a very different question of test score utilization. In this context, it is frequently important for employers to be able to ensure that each applicant selected for a position possesses some minimum of knowledge or proficiency in a subject or skill. Thus, as in the illustration presented earlier, a local school board may rely on a standardized achievement test to assure itself that applicants for teaching positions in history have some minimum knowledge of the subject to be taught.

This use of test results is distinguishable from those uses involving either overt or covert prediction. As noted,⁶² it is predictive, but only in the sense that an employer anticipates that those applicants who cannot demonstrate some level of proficiency in a subject of specialized training will not be qualified to hold down the job which is sought. Here, the employer does not make the claim that the procedure increases business efficiency by enabling him to hire only the better applicants. Rather, the justification advanced is that a core of knowledge or skills is essential for adequate execution of the tasks involved; the efficiency gain arises from avoiding the costs of employing people who are simply unable to discharge their duties—not the costs of employing less productive workers. Additionally, an employer who holds out to the public that his staff possesses some particular expertise can assert that this use of test scores is necessary to provide objective evidence in support of the claimed expertise. For example, a local school board or a private school which represents that its teachers are all competent in their respective subject areas must be able to support that claim.

These very different reasons for reliance on test results suggest that it may not be appropriate to insist on criterion validation in this special area.⁶³ Where an employer makes no claim that higher test scores

62. See text accompanying note 54 *supra*.

63. The Standards put forth by the American Psychological Association also fail to take into account this use of achievement testing: "An employer cannot justify an employment test on grounds of content validity if he cannot demonstrate that the content universe includes all, or nearly all, important parts of the job." STANDARDS 29. That statement, taken at face value, would prohibit content validation for the limited achievement testing discussed here. Support for not following the *Standards* on this point is drawn from the close analogy between this use of testing and the requirement of degrees or diplomas, see note 84 *infra* and accompanying text, and the Supreme Court's express refusal to subject such diploma requirements to the Guidelines validation procedure. *Griggs v. Duke Power Co.*, 401 U.S. 424, 433 n.8 (1971); *League of United Latin American Citizens v. City of Santa Ana*, 12 F.E.P. Cas. 651, 672 (C.D. Cal. 1976). *But*

predict better performance, and does not use the scores in a way which ranks applicants, content validation should not be presumed improper.

Justifying a Content-Valid Test. Once an employer who seeks to defend his use of an achievement test has established that test results are not utilized to rank applicants, but only to guarantee minimal knowledge or proficiency in an area of specialized training, two questions remain for the court. First, what type of showing must the employer make to overcome the Guidelines' preference for a criterion-valid test? Second, even where content validation (as opposed to criterion validation) is appropriate, in what situations may an employer make use of an achievement test which has a disproportionate racial impact?

Identifying the use made of test results clearly represents a large step toward answering the first inquiry: where applicants are ranked according to test result, criterion validation should generally be required.⁶⁴ Where ranking is not employed, however, the courts must look further to determine whether criterion validation is "feasible" under the EEOC Guidelines. This Note has urged that the Guidelines not be construed so rigidly as to require a showing of the technical impossibility of criterion validation; rather, an employer who can demonstrate the impracticability or prohibitive cost of designing a criterion-valid test should be able to resort to the less preferred technique.⁶⁵ How an employer goes about making this showing of impracticability is a question which the courts have not yet considered.

It should be recalled that an employer who seeks to use a content-valid achievement test must concede that the test results do not predict which applicants will perform better as employees. Still, the employer may claim that the achievement test does increase his business efficiency by identifying those applicants who do not possess the minimum knowledge or skill in a subject area which is important in the position. Either of two circumstances could support such an argument by the employer.

see Note, *Diplomas, Degrees, and Discrimination*, 26 HASTINGS L.J. 1377, 1388-89 (1975).

64. The only plausible justification for use of content validation in this context would be the claim that it is either impossible or prohibitively expensive to define with adequate precision the appropriate criterion of job success. The argument would be that it is not possible to rank employees as better or worse according to some performance criterion; *i.e.*, that there is no sufficiently concrete measure of job success. Simultaneously, however, the employer must assert that the factors measured by the test are so important for the job and so coincident with the total domain of job requirements that ranking is justified. Although the *Standards* apparently contemplate this situation, STANDARDS 29, it is not likely that the argument would be successful apart from jobs with very narrowly defined requisites, such as typing positions, *id.*

65. See text accompanying notes 56-57 *supra*.

First, there may be no concrete criterion of job success. That is, an employer may plausibly assert that it is either impossible or prohibitively expensive⁶⁶ to ascertain which of his employees are "better" than others.⁶⁷ Thus, while it may be relatively simple to gauge the job success of assembly line workers by measuring daily production or number of errors, it might be extremely difficult to make a quantitative assessment of the performance of an architect or a police detective.⁶⁸ At the same time, it is clear that a qualified architect must have some understanding of techniques of design and the physical properties of various building materials; the police detective must possess a working knowledge of criminal law and modern methods of crime detection. Content-valid tests should therefore be permitted in this context.

The second circumstance in which content validation should be allowed may best be illustrated by reviewing the earlier example of the high school history teacher.⁶⁹ Assuming that a success criterion for history teachers is ascertainable,⁷⁰ one might still show the infeasibility of using a criterion-valid test by demonstrating that some of the qualities which make a good teacher are not measurable by current testing techniques. A properly conducted job analysis might reveal, for example, that a high school history teacher needs qualities A, B, C, D, E, and F. Of these qualities, only A, B, and C—a knowledge of the subject matter taught in history classes, proficiency in pedagogical techniques, and an understanding of the principles of psychology—can be measured

66. "Prohibitively expensive" would mean that the cost of designing and implementing a criterion-valid testing process would exceed the value of efficiency gains to be derived from use of the test results. That is, it may be cheaper for an employer to hire some less qualified employees and "weed them out" over time than to screen out that group in advance. It should be noted that there may be a substantial difference between the cost of employing "less competent" workers and the cost of hiring "incompetent" workers.

67. It may be extremely difficult to define an appropriate criterion of job success, see note 68 *infra* and accompanying text, and it is obviously crucial to the adequacy of the test that the criterion itself be valid. STANDARDS 27. There is some evidence, for example, that an employee's sex and race tend to bias the job performance ratings given by supervisors. Bigoness, *Effect of Applicant's Sex, Race, and Performance on Employees' Performance Ratings: Some Additional Findings*, 61 J. APPLIED PSYCHOLOGY 80 (1976).

68. Cf. Hughes, *Mistakes at Work*, 17 CANADIAN J. ECON. & POL. SCI. 322 (1951), reprinted in PROFESSIONALIZATION 148 (H. Vollmer & D. Mills eds. 1966).

69. See notes 25-29 *supra* and accompanying text.

70. One possible criterion would be the performance of those students taught by the instructor in a content-valid achievement test for history. Another would be the evaluation of an instructor by his colleagues. Cf. Wilson, Dienst & Watson, *supra* note 27; Subkoviak & Levin, *supra* note 27. The deficiencies inherent in both of these criteria are apparent. See note 67 *supra*.

by tests which are not prohibitively expensive. Factors D, E, and F—"patience, warmth [and] humor"⁷¹—cannot be measured in advance by a test. Because several crucial factors are not measurable, it would not be reasonable for a local school board to attempt to design a test which would be used to rank applicants. Yet the importance of those factors which are measurable strongly suggests that the school board should be permitted to test for some minimal proficiency in these areas.⁷²

In these two situations⁷³—where there is no appropriate criterion of job success or where there are factors which are immeasurable that may contribute to success—the courts should adopt a balancing approach which is structurally similar to the balancing mandated for hiring decisions by the Supreme Court in *Griggs v. Duke Power Co.*,⁷⁴ but with an emphasis on different factors. *Griggs* requires the balancing of a racially disproportionate impact against the increased efficiency which results from use of the challenged test. In parallel fashion, a decision to permit a content-valid test where use of a criterion-valid test would be impractical should be based on a weighing of the reduction of disproportionate impact which a criterion-valid test would yield against the added expense and difficulty of designing and implementing such a test.⁷⁵ The crucial question, therefore, is the extent to which criterion validation improves upon content validation in removing the discriminatory impact of the test.

The Burden of Proof. Allocation of the burden of proof on this question is of fundamental importance. If the employer must demonstrate that a change to a criterion-valid test would be ineffective in reducing the disproportionate impact, he is unlikely to be able to meet

71. See note 28 *supra* and accompanying text.

72. The cut-off point itself must be valid if the testing procedure is to pass muster. An arbitrary cut-off, see *United States v. North Carolina*, 400 F. Supp. 343, 349 (E.D.N.C. 1975) (three-judge court), or one set simply to limit the influx of potential employees, see *Walston v. County School Bd.*, 492 F.2d 919, 925 (4th Cir. 1974), should not be permitted if the testing procedure has a disproportionate racial impact.

73. The two situations described are not mutually exclusive. It is possible, indeed likely, that a position for which no appropriate success criterion can be established also requires several employee skills or traits which are not measurable by a pre-employment objective testing device. See *Sandman & Urban*, *supra* note 32, at 48.

74. 401 U.S. 424 (1971). See Note, *supra* note 10, at 101. See also Affeldt, *Title VII in the Federal Courts—Private or Public Law*, 15 VILL. L. REV. 1, 31 (1969); Wilson, *supra* note 21, at 851; Note, *Application of the EEOC Guidelines to Employment Test Validation: A Uniform Standard for Both Public and Private Employers*, 41 GEO. WASH. L. REV. 505, 508, 521 (1973).

75. It is important to note that the question considered here is not whether the testing device may be used at all, see notes 41-54 *supra* and accompanying text, but whether a criterion-valid procedure is a reasonable alternative to the challenged content-valid test. Cf. Note, *supra* note 10.

his burden except where he can show the virtual impossibility of designing a criterion-valid procedure. That is, the employer would be placed in the position of having to prove the negative proposition that a different type of test would *not* lead to any significant improvement. Such a showing clearly would be most difficult and very expensive, and it is unlikely that employers would be able—or willing—to make it.⁷⁶ As a consequence, a court which balances the potential reduction of disproportionate impact against the difficulty of implementing a criterion-valid procedure will inevitably find for the plaintiff unless it can be shown that the preferred test is technically impossible to design.

Requiring the employer to show that there is no less discriminatory alternative to his testing procedure is therefore plainly unworkable.⁷⁷ Rather, the *plaintiff* should be required to prove the converse proposition—that there is a criterion-valid procedure which would reduce the undesirable impact of a content-valid test.⁷⁸ In the first situation described above,⁷⁹ where the employer asserts that there is no appropriate criterion of job success, a plaintiff should be required to show two things before a content-valid test can be struck down simply because it is not criterion-valid. First, it must be shown that it is neither impossible nor prohibitively expensive to fashion a criterion of job success.⁸⁰ Second, there should be a showing of some probability that such a criterion-valid procedure would indeed have a less disproportionate impact. Similarly, in the second situation, where the criteria of success are only partially ascertainable but where several important factors are measura-

76. Indeed, many employers have already abandoned the use of objective tests rather than incur the expense of validation and litigation. See note 7 *supra* and accompanying text.

77. See Note, *supra* note 10, at 113.

78. *Id.* The Supreme Court apparently has adopted this approach:

If an employer does then meet the burden of proving that its tests are "job related," it remains open to the complaining party to show that other tests or selection devices, without a similarly undesirable racial effect, would also serve the employer's legitimate interest in "efficient and trustworthy workmanship." *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975) (citation omitted).

79. See notes 66-68 *supra* and accompanying text.

80. Consider, for example, *Coopersmith v. Roudebush*, 517 F.2d 818 (D.C. Cir. 1975), wherein an applicant for a position as an attorney with the Board of Veterans Appeals charged sex discrimination in a test given for that position. The Board had attempted to fashion a criterion of job success in the form of performance on a typical case handled by the Board attorneys, and pre-tested applicants by requiring them to work through the sample problem. (The charges stemmed from the subjective evaluation which must accompany such a loosely structured test.)

Other methods a plaintiff might use to establish the appropriateness of a success criterion include evidence of the practices of similar employers and testimony by job analysis experts.

ble,⁸¹ the plaintiff should be required to demonstrate both that there are means of measuring the factors not tested for and that there is at least a reasonable probability that the addition of these factors to the test would reduce its disproportionate impact.

It is clear that such proof will be difficult for the plaintiff to marshal. Moreover, where it is practicable to use a criterion-valid screening mechanism, most employers who seek to maximize efficiency in their selection procedures will already have adopted the preferred technique voluntarily. An employer is likely to be reluctant to change testing procedures only where the increase in efficiency which he estimates is less than the cost of designing and implementing the new test.⁸² In such a case, the role of a reviewing court would be to evaluate the legitimacy of the balancing already performed by the employer.

This apparently heavy burden on the plaintiff may seem unwarranted in light of the Title VII mandate that discriminatory testing is not permissible unless a business justification is present. It must be remembered, however, that the plaintiff's burden is not to prove the absence of a business justification, but to demonstrate the feasibility of criterion validation once the employer has shown that his use of test results is an appropriate content-valid procedure. Even where the plaintiff fails to carry this burden, the employer is still required to show a business justification for the use of the test.

Business Justification. A determination of business justification depends on an evaluation of the "job-relatedness" of the test.⁸³ In making this determination, a reviewing court should ask the same questions which are asked in scrutinizing diploma and degree requirements.⁸⁴ Use of a content-valid test is analytically similar to a degree

81. See notes 69-72 *supra* and accompanying text.

82. There is, of course, the possibility that the employer may be using the test results as a covert form of intentional racial discrimination or as a pretext for racially motivated hiring. See *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 426 (1975); *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 804-05 (1973). Where there is evidence of such covert discrimination, the court should require a very convincing demonstration by the employer of the business efficiency to be derived from the testing procedure. Cf. *Walston v. County School Bd.*, 492 F.2d 919, 924-25 (4th Cir. 1974); *Baker v. Columbus Municipal Separate School Dist.*, 462 F.2d 1112, 1114 (5th Cir. 1972).

83. See *Griggs v. Duke Power Co.*, 401 U.S. 424, 431-32 (1971). See also *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 424 (1975); *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 802-03 (1973); *Sandman & Urban*, *supra* note 32, at 38-39; Note, *supra* note 10, at 107-13.

84. See, e.g., *Payne v. Travenol Laboratories*, 12 F.E.P. Cas. 770 (N.D. Miss. 1976); *League of United Latin American Citizens v. City of Santa Ana*, 12 F.E.P. Cas. 651 (C.D. Cal. 1976). For a criticism of the use of diploma and degree requirements,

requirement in that, while in neither case are all the factors contributing to success on the job measured, both a passing test score and the possession of a degree are believed to demonstrate proficiency in a subject or skill which is important to job performance.⁸⁵ The court should therefore ask: (1) whether the test measures only those skills or subjects important in the job,⁸⁶ and (2) whether the subjects measured are significant in relation to the total requirements of the position. The answers to these questions will guide the court in determining how much efficiency the employer will gain by using the content-valid test. That determination will be balanced according to the *Griggs* standard against the magnitude of the disproportionate impact to decide whether the testing procedure is permissible.⁸⁷

If achievement testing for jobs which require specialized training is not to be discouraged unnecessarily, another factor should be taken into account in this balancing process. A court should give some deference to the steps taken by an employer in the preparation of testing procedures. Where a test is prepared or selected haphazardly, for instance, by a personnel director who makes the selection without the aid of psychological testing experts, the courts should require a very strong showing of increased efficiency. A test prepared at some expense by recognized experts in testing, on the other hand, should be accorded some presumption of validity. An employer using a test in the latter category should not be forced to bear the great expense of duplicating the validation of his testing procedures where both the test and the appropriate cut-off have been determined in a systematic way. This "sliding scale" approach to test scrutiny has been used in the Second Circuit,⁸⁸ and would seem to be appropriate for application in all Title

and the argument that such selection devices should be subject to the EEOC Guidelines, see Note, *supra* note 63. See also Huff, *supra* note 28.

85. It can be argued that the achievement test device is a significant improvement over degrees and diplomas, since the former is not complicated by variations in quality among schools. In *League of United Latin American Citizens v. City of Santa Ana*, 12 F.E.P. Cas. 651 (C.D. Cal. 1976), for example, there was evidence that the high schools in the Santa Ana area would graduate any student who could read at a fifth-grade level. *Id.* at 657.

86. This question is generally answered by a job analysis, but it raises very difficult and important issues where specialized training is required for a job. See notes 37-38 *supra* and accompanying text.

87. Where there is a history of overt discrimination by the employer, the court will require a very strong showing of efficiency gains through the "job-relatedness" of the test. See note 99 *infra*.

88. *Vulcan Soc'y of N.Y. City Fire Dep't, Inc. v. Civil Serv. Comm'n*, 490 F.2d 387, 395-96 (2d Cir. 1973). Two recent applications of the approach are found in *Jones v. New York City Human Resources Admin.*, 528 F.2d 697, 699 (2d Cir. 1976), and

VII cases.⁸⁰

THE FOURTEENTH AMENDMENT STANDARD AND ITS
CONTINUING VIABILITY IN THE REVIEW OF TESTING

The Traditional Standard of Review

The fourteenth amendment is used to challenge tests in circumstances in which Title VII does not apply. Although the scope of Title VII has not yet been conclusively determined,⁹⁰ it appears that state certification procedures which are designed to determine whether an individual possesses the expertise needed in a particular field are subject only to the more relaxed standards of fourteenth amendment review. The traditional fourteenth amendment standard was articulated in *Dent v. West Virginia*.⁹¹ The Court stated that there is no deprivation of a right guaranteed by the fourteenth amendment when a state refuses to license an individual due to his failure to comply with regulations promulgated for the general welfare.⁹² It is within the power of the state, the Court concluded, "to prescribe all such regulations as, in its judgment, will secure or tend to secure [its people] against the conse-

Kirkland v. New York State Dep't of Correctional Serv., 520 F.2d 420, 425-26 (2d Cir. 1975).

89. A three-judge federal court in North Carolina may have taken the same approach toward the validation of a cut-off score on a standardized teaching examination. See *United States v. North Carolina*, 400 F. Supp. 343, 348-49 & n.6 (E.D.N.C. 1975). Carried to its logical conclusion, an absolute refusal to permit content validation of professional competence examinations would necessitate a complete restructuring of professional recruitment practices. Consider the example of the legal profession. The vulnerability of the bar examination is obvious: the impossibility of defining "success" as a lawyer makes it equally impossible to determine whether a given test is an adequate predictor of success. Beyond this, however, the traditional reliance of legal employers on law school grades as a selection criterion would also be subject to attack. An employer using law school grades for selection purposes is implicitly relying on law school examinations as valid ranking tools. It is doubtful, though, whether it could be shown that law school course examinations are accurate measures of important job skills, or even that law school courses teach such skills. For a discussion of these issues in the recently controversial context of the North Carolina teacher certification process, see *Charlotte Observer*, July 11, 1976, § B, at 2, col. 1. See notes 100-02 *infra* and accompanying text.

90. See note 16 *supra*.

91. 129 U.S. 114 (1889).

92. *Id.* at 122. States have long used educational requirements and proficiency examinations to ascertain whether an individual is qualified to practice in a field, such as medicine or law, which directly affects the general welfare of its people.

[I]t has been the practice of different States from time immemorial, to exact in many pursuits a certain degree of skill and learning upon which the community may confidently rely, their possession being generally ascertained upon an examination of parties by competent persons *Id.*

quences of ignorance and incapacity as well as of deception and fraud."⁹³ In sustaining a West Virginia statute which required a diploma from a recognized institution or an adequate performance on an examination for certification as a physician, the Court held that such a requirement could only be overturned if it bore no reasonable relation to the practice of the profession: "If [the qualifications] are appropriate to the calling or profession, and attainable by reasonable study or application, no objection to their validity can be raised because of their stringency or difficulty."⁹⁴

The Title VII showing of "racially disproportionate impact" has not been acknowledged by the Supreme Court as sufficient to establish a prima facie case of deprivation of fourteenth amendment rights.⁹⁵ The traditional constitutional standard was most recently reaffirmed in *Washington v. Davis*, an employment testing challenge brought under the fifth amendment.⁹⁶ The Court specifically rejected an effort to

93. *Id.* Such requirements are time-tested means of assuring that practitioners are competent. *Rasulius v. Weinberger*, 502 F.2d 1006, 1010 (7th Cir. 1974) (involving the licensing of physical therapists). See also *Western Addition Organization v. Alioto*, 360 F. Supp. 733, 734 (N.D. Cal. 1973) (test for policemen struck down, but court noted that city had the right to test); *England v. Louisiana Bd. of Medical Examiners*, 246 F. Supp. 993 (E.D. La. 1965), *aff'd mem.*, 384 U.S. 885 (1966) (not irrational and unreasonable for state legislature to require chiropractors to comply with licensing provisions).

94. 129 U.S. at 122. As long as they apply uniformly to all who seek certification, such qualifications will not be deemed arbitrary. See *id.* at 124.

95. Even where it has had the opportunity to do so, the Supreme Court has refused to hold conclusively that a policy that is facially neutral and evenly applied is violative of the fourteenth amendment because of its disproportionate impact. *Tyler v. Vickery*, 517 F.2d 1089, 1097-98 (5th Cir. 1975), *cert. denied*, 44 U.S.L.W. 3719 (U.S. June 14, 1976), *citing* *Geduldig v. Aiello*, 417 U.S. 484 (1974); *cf.* *James v. Valtierra*, 402 U.S. 137, 141 (1971) (upholding amendment to state constitution requiring a referendum for low-income housing).

There are, however, a number of circuit court decisions which appear to apply the disproportionate impact test to purely constitutional challenges to employment testing procedures. See note 103 *infra* and accompanying text. Such cases have been strongly criticized. *Tyler v. Vickery*, 517 F.2d at 1095-97.

96. 44 U.S.L.W. 4789 (U.S. June 7, 1976). In *Davis*, two blacks whose applications to become District of Columbia police officers had been rejected challenged a written test of verbal aptitude which had been used in the selection process and "which excluded a disproportionately high number of Negro applicants." *Id.* at 4790. Finding the test to be neutral on its face and rationally related to the legitimate government purpose of upgrading the verbal abilities of police officers, the Court held the "more vigorous" Title VII standard inapplicable:

A rule that a statute designed to serve neutral ends is nevertheless invalid, absent compelling justification, if in practice it benefits or burdens one race more than another would be far reaching and would raise serious questions about, and perhaps invalidate, a whole range of tax, welfare, public service, regulatory, and licensing statutes that may be more burdensome to the poor and to the average black than to the more affluent white.

. . . [E]xtension of the rule beyond those areas where it is already ap-

extend the applicability of Title VII to non-statutory contexts, and required instead a showing of discriminatory purpose. The significance of this decision in the testing area is apparent. If it is assumed that racial differences in test performance reflect cultural and educational variations, and are not based on disparities in genetically determined cognitive abilities,⁹⁷ then it will be virtually impossible to prove that a test is racially discriminatory in itself—the rejoinder will be that the test itself is not biased, but is merely an accurate reflection of underlying social conditions.⁹⁸ Therefore, claiming that a test is racially biased will be feasible only if the test is clearly discriminatory or where there is evidence that it is being used in an intentionally discriminatory manner.⁹⁹

plicable by reason of statute, such as in the field of public employment, should await legislative prescription. *Id.* at 4794.

97. See *Arrington v. Massachusetts Bay Transit Authority*, 306 F. Supp. 1355, 1358 (D. Mass. 1969), quoted in note 8 *supra*; Note, *Employment Discrimination: Statistics and Preferences Under Title VII*, 59 VA. L. REV. 463, 474 (1973).

98. It is for this reason that recent efforts to require differential validation by race will be inadequate to correct many of the most dangerous effects of employment testing. As a result of the numerous studies emphasizing the need for validation by race, see, e.g., Rudy & Albright, *Racial Differences on Selection Instruments Related to Subsequent Job Performance*, 21 PERSONNEL PSYCH. 31 (1968), both the EEOC, see 29 C.F.R. § 1607.4 (1975), and the courts, see *Sandman & Urban*, *supra* note 32, at 53, have begun to insist on differential validation with increasing frequency.

Validation by race solves only part of the problem. As the authors of one psychological study recently concluded:

[Those] concerned with the applicability of employment tests to minority groups should probably direct their future efforts to the study and determination of test fairness rather than to the pseudoproblem of racial differences in [test validity]. Schmidt, Berner & Hunter, *Racial Differences in Validity of Employment Tests: Reality or Illusion?*, 58 J. APPLIED PSYCH. 5, 8 (1973).

See also *Arrington v. Massachusetts Bay Transit Authority*, 306 F. Supp. 1355, 1358 (D. Mass. 1969), quoted in note 8 *supra*. Modern analytic methods have apparently begun to reduce the racial disparities in testing validity. See Campbell, *Tests Are Valid for Minority Groups Too*, 2 PUB. PERSONNEL MANAGEMENT 70 (1973) (concluding that aptitude tests predict as well for minorities and that where the test is apt to be biased, it is more likely to overpredict job performance for minority group members). The problem which must now be addressed concerns the actual fairness of the whole testing procedure.

99. A history of racial discrimination may provide support for such a challenge. Many of the employment testing cases arise in factual situations where discrimination was openly practiced in the past, see, e.g., *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975); *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), and evidence of such practices might be used to support an inference of discrimination. Particularly where the validity of the test is dubious, a history of discrimination might tilt the scales so that a court would be forced to conclude that the only possible reason for establishing testing procedures was discriminatory. A testing procedure with such an impermissible purpose would have to fall under a rational basis test. See note 82 *supra* and accompanying text.

Indeed, it should be emphasized that the rational basis test can be used to challenge a testing procedure in any case where the test is so invalid that it has no relationship to

This traditional approach to the review of allegedly discriminatory testing procedures was also recently employed by the Fifth Circuit to uphold the Georgia bar examination despite the absence of any validation study. In *Tyler v. Vickery*,¹⁰⁰ a class action was brought on behalf of all blacks who had taken and failed the state bar examination. The test was challenged as unconstitutional on due process and equal protection grounds. Although black applicants as a class had experienced great difficulty in passing the bar examination,¹⁰¹ the court held the test to be a rational means for the state to exclude persons who do not meet its standards of competence.¹⁰²

The Future of the Traditional Standard

The traditional fourteenth amendment standard of review is being questioned by the courts. The reasoning in *Griggs* has been extended from cases brought under Title VII to those brought only on constitutional grounds; many decisions ostensibly made on fourteenth amendment grounds by lower courts have blurred the distinction between the two standards.¹⁰³

what it seeks to predict, since such a test would fall under the more lenient arbitrary and capricious test of economic equal protection.

100. 517 F.2d 1089 (5th Cir. 1975), *cert. denied*, 44 U.S.L.W. 3719 (U.S. June 14, 1976).

101. In July 1972, all of the black applicants failed the examination. In the two succeeding examinations, more than one-half of the black applicants did not pass, as compared with approximately one-third of the white applicants. *Id.* at 1092.

102. The court found that the bar examination had a rational relationship to an examinee's capacity to enter the profession of law. *Id.* at 1103.

The courts have not been critical in their review of bar examinations. *See, e.g., Staley v. State Bar*, 17 Cal. 2d 119, 121, 109 P.2d 667 (1941). In *Feldman v. State Bd. of Law Examiners*, 438 F.2d 699 (8th Cir. 1971), the Eighth Circuit applied a rational basis test to uphold an Arkansas bar examination in the face of an applicant's claim of a violation of constitutional rights:

[It cannot] be said that an essay type examination is inherently unfair or that such a test has no rational connection with an applicant's fitness or capacity to practice law. *Id.* at 705.

See also Whitfield v. Illinois Bd. of Law Examiners, 504 F.2d 474 (7th Cir. 1974).

103. *See, e.g., Walston v. County School Bd.*, 492 F.2d 919 (4th Cir. 1974); *Chance v. Board of Examiners*, 458 F.2d 1167 (2d Cir. 1972); *Carter v. Gallagher*, 452 F.2d 315 (8th Cir. 1971), *cert. denied*, 406 U.S. 950 (1972). *See generally Comment, Equal Protection and Standardized Testing*, 44 *Miss. L.J.* 900, 925-30 (1973).

This blending is particularly evident in cases involving the employment of teachers. As originally enacted, Title VII contained a specific exemption for states and other governmental employers. Civil Rights Act of 1964, Title VII, Equal Employment Opportunity, Pub. L. No. 88-352, § 701(b), 78 Stat. 253, *as amended*, 42 U.S.C. § 2000e(b) (Supp. III 1973). The courts apparently compensated for this exemption by subjecting school boards to the same level of scrutiny that would be required under the statutory standard on the premise that private concerns should not be subject to more

If it is assumed that the EEOC possesses expertise in the field of validating employment tests, then it is arguable that the courts should measure tests against the Guidelines, even in situations where they are not technically applicable.¹⁰⁴ In *Washington v. Davis*,¹⁰⁵ however, the Court declined to accept the argument found persuasive in the court of appeals¹⁰⁶ that the Guidelines should be used to give substance to the fourteenth amendment in the employment area. The Fifth Circuit had previously reached a similar conclusion in a fourteenth amendment context in *Tyler v. Vickery*.¹⁰⁷

Other pre-*Davis* lower court decisions were premised on a belief that equal protection jurisprudence must take cognizance of developing psychological knowledge. A three-judge district court in the Fourth Circuit, for example, relied upon psychology experts in reaching its decision in a fourteenth amendment suit brought by blacks charging invidious discrimination due to a teacher certification examination. The court in *United States v. North Carolina*,¹⁰⁸ while recognizing the state's interest in securing competent educators, refused to allow the state legislature to set a minimum score for licensing teachers without a validation study to prove that the cut-off point was related to teaching qualifications.¹⁰⁹ By not automatically sanctioning the test, the court departed from the traditional standard.

The creation of the EEOC, with its mandate to formulate Guide-

rigorous standards than public employers. *Tyler v. Vickery*, 517 F.2d 1089, 1095-97 (5th Cir. 1975), *cert. denied*, 44 U.S.L.W. 3719 (U.S. June 14, 1976); *see Wilson, supra* note 21, at 848-51; Comment, *supra*, at 910-22.

The intentional blending in this area is no longer necessary because the 1972 amendments to Title VII deleted the exemption for governmental units. *See Grossman, supra* note 21, at 370-71.

104. *See* Comment, *supra* note 21, at 920.

105. 44 U.S.L.W. 4789 (U.S. June 7, 1976).

106. *Davis v. Washington*, 512 F.2d 956 (D.C. Cir. 1975).

107. 517 F.2d 1089 (5th Cir. 1975), *cert. denied*, 44 U.S.L.W. 3719 (U.S. June 14, 1976). The Fifth Circuit found *Geduldig v. Aiello*, 417 U.S. 484 (1974), to be persuasive authority. In *Geduldig*, it was claimed that a state-administered compensation plan which excluded from coverage disabilities associated with normal pregnancy was sex discrimination in violation of the equal protection clause of the fourteenth amendment. The Supreme Court upheld the state plan by applying the relaxed rational relationship standard of review and ignored an EEOC Guideline, 29 C.F.R. § 1604.10(b) (1975), which took a contrary position. This case was analogized by the *Tyler* court to the racial discrimination issue.

Note that this position does have the advantage of leaving open the issue of content validation, rather than accepting the Guidelines as rigid standards.

108. 400 F. Supp. 343 (E.D.N.C. 1975) (three-judge court).

109. Cut-off scores can be validated. The designers of the test given in North Carolina approved higher minimum scores for use in Chicago and Los Angeles. *Id.* at 349 n.6.

lines for equal employment, might be interpreted as a general expression of congressional intent that contemporary psychological expertise should be incorporated into equal protection analysis of employment practices.¹¹⁰ The obvious counterpoint is that Congress intentionally limited Title VII's scope of application.¹¹¹ On balance, however, it seems incongruous that Congress would incorporate modern scientific theory into the vast majority of employment-related cases while leaving the remainder to be decided on classical philosophical grounds.

The most effective approach might be to make use of the Second Circuit's "sliding scale" analysis¹¹² in fourteenth amendment employment cases. If a test has been designed in accordance with professionally accepted practices, then scrutiny should be limited to the traditional fourteenth amendment reasonableness test. Where the test in question has been prepared or selected arbitrarily, however, the courts should impose a standard of review which approximates that used in Title VII actions. In this way, the courts can give all litigants the benefit of evolving principles of equal protection in the employment area without undertaking unwarranted extension of a statute which is expressly limited in its applicability.

CONCLUSION

The original purpose of employment testing was to promote merit hiring and thereby eliminate biases not related to qualifications. Ironically, many of the professionally developed tests now being struck down by the courts were intended to counteract the prejudices of those selecting among applicants.¹¹³ If tests do serve a protective function,¹¹⁴ then the gains for qualified members of minority groups will be small if the courts so discourage testing that reversion to subjective selection techniques becomes widespread.¹¹⁵ The ultimate result of the current trend under the EEOC Guidelines appears to be quota hiring¹¹⁶ undertaken to

110. See note 42 *supra*.

111. See note 16 *supra*.

112. See note 88 *supra* and accompanying text.

113. See *Boston Chapter, NAACP, Inc. v. Beecher*, 504 F.2d 1017, 1022 (1st Cir. 1974). It has been recognized that the stringent standards are making validation of tests very difficult. See *Developments in the Law, supra* note 7, at 1122.

114. See *Cooper & Sobol, supra* note 7, at 1664 n.124.

115. These subjective techniques are obviously more difficult to control for discrimination. See *Comment, supra* note 21, at 924.

116. See *Wilson, supra* note 21, at 873-74. This in itself is offensive to the theory underlying Title VII. Note, *supra* note 97, at 466 n.15. Justice Blackmun, in his concurring opinion in *Moody*, wrote that

too-rigid application of the EEOC Guidelines will leave the employer little choice, save an impossibly expensive and complex validation study, but to en-

avoid litigation over subjective techniques,¹¹⁷ despite the express rejection of quotas in Title VII.¹¹⁸

A middle ground is needed between the relaxed fourteenth amendment standard and the stringent EEOC Guidelines, especially for evaluating tests used to determine whether a person possesses a required minimum level of knowledge or skill. In this highly technical field, the agencies and the courts should make greater use of the expertise of professional psychologists and accept content validation as appropriate for tests measuring achievement.

gage in a subjective quota system of employment selection. This, of course, is far from the intent of Title VII. 422 U.S. at 449.

117. Subjective techniques will also be required to meet EEOC standards:

Selection techniques other than tests, as defined in § 1607.2, may be improperly used so as to have the effect of discriminating against minority groups. Such techniques include, but are not restricted to, unscored or casual interviews and unscored application forms. Where there are data suggesting employment discrimination, the person may be called upon to present evidence concerning the validity of his unscored procedures as well as of any tests which may be used, the evidence of validity being of the same types referred to in §§ 1607.4 and 1607.5. . . . 29 C.F.R. § 1607.13 (1975).

118. See note 5 *supra*.