Title VII and Competitive Testing

James M. Conway
NOTES

TITLE VII AND COMPETITIVE TESTING

As of 1984, federal, state, and local government employees numbered over sixteen million, with a combined payroll of nearly twenty-seven billion dollars. This vast employment network has relied on competitive exams to ascertain the merit and fitness of potential employees, and they are the determining factors for most civil service positions.


Employees (1,000) in 1984:

State .............................................................................. 3,898
Local .............................................................................. 9,595
Federal (civilian) ............................................................. 2,942

October 1984 Payrolls (million dollars):

State .............................................................................. 5,815
Local .............................................................................. 13,952
Federal (civilian) ............................................................. 7,137

Id.

service positions. The use of competitive exams has decreased due to recent charges of discriminatory effect, and many exams have been invalidated pursuant to Title VII of the Civil Rights Act of 1964.

Elimination of discriminatory preferences and the establishment of fair employment practices are the goals of Title VII, and these goals must precede the legislative goals advanced by merit testing. Title VII was neither intended to guarantee employment to members of any minority group, nor was it intended to preclude the use of exams to determine job qualifications. Title VII has been applied beyond its intended scope and is effectively eliminating the use of competitive physical exams for those civil service positions requiring


5. See infra note 13.

6. Griggs v. Duke Power Co., 401 U.S. 424, 431 (1971) (requiring "the removal of artificial, arbitrary, and unnecessary barriers to employment . . ."). In Griggs, the Supreme Court expanded the scope of possible liability under Title VII by holding that employment practices that have a disparate impact on a protected class, although not discriminating on their face, are unlawful unless justified by "business necessity." Id. at 431.

7. Guardians Ass'n of New York v. Civil Serv. Comm'n, 630 F.2d 79, 104-05 (2d Cir. 1980), cert. denied, 452 U.S. 940 (1981). Title VII "explicitly relieves employers from any duty to observe a state hiring provision 'which purports to require or permit' any discriminatory employment practice." Id. at 105 (quoting 42 U.S.C. § 2000e-7 (1976)).

8. Griggs, 401 U.S. at 430 (stating that "Congress did not intend by Title VII . . . to guarantee a job to every person regardless of qualifications.").

9. Id. at 436.
physical fitness.

This Note analyzes the application of Title VII to competitive physical exams and the deterioration of defenses available to the civil service employer. This Note concludes that civil service tests cannot continue to be administered in the present form, which includes an automatic Title VII challenge. The aggregate economic cost and the reduction in the quality of municipal services are major consequences of the seemingly endless Title VII litigation and must not be dismissed as unavoidable side effects of Title VII.

I. TITLE VII

Title VII prohibits employment practices which discriminate on the basis of an "individual's race, color, religion, sex, or national origin," in an attempt to remove "artificial, arbitrary, and unnecessary" employment practices which result in discrimination. The Congressional commitment to eliminate invidious discrimination is illustrated by the liberal standards for establishing a Title VII violation and the inclusion of the backpay provision. The backpay provision of Title VII, which is modeled on the corresponding provision

10. See infra notes 85-88 and accompanying text. Written tests, which are ideally suited to content validation, are more likely to withstand Title VII scrutiny. See, e.g., Craig v. County of Los Angeles, 626 F.2d 659, 664, 668 (9th Cir. 1980)(written exam was sufficiently job-related, but height requirement was invalid), cert. denied, 450 U.S. 919 (1981); Allen v. City of Mobile, 466 F.2d 122 (5th Cir. 1972)(police written exam was job-related), cert. denied, 412 U.S. 909 (1973).

11. See infra notes 37-65 and accompanying text.


(a) Employer practices

   It shall be unlawful employment practice for an employer—

   (1) to fail or refuse to hire or to 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; or

   (2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual's race, color, religion, sex, or national origin.

Id.


of the National Labor Relations Act,\textsuperscript{17} is intended to compel employers to examine hiring policies with greater scrutiny and to eliminate any discriminatory practices, while also compensating for any injury suffered as a result of employment discrimination.\textsuperscript{18}

Establishment of a prima facie case of discrimination only requires a discriminatory effect,\textsuperscript{19} which exists when a group protected by Title VII is adversely affected by an employment policy.\textsuperscript{20} Intent to discriminate is not required to establish a Title VII "disparate impact" case,\textsuperscript{21} but would be necessary in an alternate Title VII cause of action based on "disparate treatment."\textsuperscript{22}

\textsuperscript{17} Id. at 419 n.11.
\textsuperscript{18} Id. at 417-19. See also United States v. N.L. Industries, Inc., 479 F.2d 354, 379 (8th Cir. 1973)(backpay induces employers to correct errant employment procedures without court intervention). See generally Developments in the Law, Employment Discrimination And Title VII Of The Civil Rights Act Of 1964, 84 Harv. L. Rev. 1109, 1163-64, 1242-1269 (1971)(a detailed discussion of the remedies available under Title VII and their underlying policies).
\textsuperscript{22} International Bhd. of Teamsters v. United States, 431 U.S. 324, 335 n.15 (1977). In disparate treatment cases, the employer treats some individuals "less favorably than others because of their race, color, religion, sex, or national origin." Id. In these situations "[p]roof of discriminatory motive is critical." Id. In disparate impact cases, employment practices which are facially neutral, have a discriminatory effect on some protected group and proof of discriminatory motive is not required. Id.

Burdens of proof differ with these alternate causes of action under Title VII. See Texas Dep't of Community Affairs v. Burdine, 450 U.S. 248, 252-53 (1981). Disparate treatment involves a claim of discrimination by an individual and the plaintiff must show that he or she is a member of a protected group, who applied and was qualified for an available position, and was rejected under circumstances that raise an inference of unlawful discrimination. Id. at 253-54; McDonnell Douglas Corp. v. Green, 411 U.S. 792, 802 n.14 (1973). In a disparate impact case, the "plaintiff claims that a facially neutral employment policy has a discriminatory impact on protected classes." Burdine, 450 U.S. at 252 n.5. See also Furnish, A Path Through The Maze: Disparate Impact And Disparate Treatment Under Title VII Of The Civil Rights Act Of 1964 After Beazer And Burdine, 23 B.C.L. Rev. 419, 436-40 (1982)(a clear theoretical delineation between these two violations, detailing the burdens of proof and the available defenses, has not been articulated).

The disparate impact must be effectuated upon a group. A court should not apply a disparate impact test in a situation involving an individual. The individual protection goal of Title VII must not be confused with the legal theories designed to achieve these goals. Connecticut v. Teal, 457 U.S. 440, 458-59 (1982)(Powell, J., dissenting).

Title VII also does not preclude an individual from asserting a constitutional claim of
"[L]iberal substantive standards . . . , including the usefulness of statistical proof," are sufficient to establish a Title VII disparate impact violation. The statistical showing does not have to be based on actual applicants, but can be based on general population statistics. While evidence describing how the disparate impact was effectuated is not necessary, the selection rate of the protected group must be less than eighty percent of the selection rate of a comparison group to establish disparate impact.

Although Title VII does not require a discriminatory motive or purpose, other restrictions separate it from § 1983 violations. Section 1983 is subject to state statutes of limitation, since these violations are characterized as personal injury actions. Wilson v. Garcia, 471 U.S. 261, 265-66 (1985). Title VII violations, however, must be brought within 180 days of the alleged practice. 42 U.S.C. § 2000e-5(e) (1982). Section 1983 violations also allow compensatory, as well as punitive damages in certain circumstances. Carey v. Piphus, 435 U.S. 247, 257 n.11 (1978) (for the purpose of deterring or punishing violations of constitutional rights). Furthermore, Title VII's legal relief may include reinstatement with or without backpay. 42 U.S.C. § 745, 754 (1966)(requiring state action to trigger the equal protection clause of the fourteenth amendment, disproportionate impact is not sufficient. Washington v. Davis, 426 U.S. 229, 239 (1976). Discriminatory purpose implies that a particular course of action was selected or reaffirmed because of its adverse impact on a specific group. Personnel Adm'r of Mass. v. Feeney, 442 U.S. 256, 279 (1979).


Adverse impact and the "four-fifths rule." A selection rate for any race, sex, or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for
A minority of circuit courts limit the disparate impact analysis to specific employment practices, and require that the aggrieved plaintiff identify the particular offending employment practice which exists in a multicomponent employment system. The disparate impact analysis is therefore not applicable to multicomponent employment systems. This limitation is an attempt to increase the burden on the plaintiff. Proof of disparity alone would not be sufficient to establish a prima facie case; the plaintiff would be required to identify the employment practice which caused the discrimination.

The group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact. Smaller differences in selection rate may nevertheless constitute adverse impact, where they are significant in both statistical and practical terms or where a user's actions have discouraged applicants disproportionately on grounds of race, sex, or ethnic group.

27. Pouncy v. Prudential Ins. Co., 668 F.2d 795, 800 (5th Cir. 1982); American Fed'n of State, County & Mun. Employees v. Washington, 770 F.2d 1401, 1405 (9th Cir. 1985); Robinson v. Polaroid Corp., 732 F.2d 1010, 1014-17 (1st Cir. 1984); Mortensen v. Callaway, 672 F.2d 822, 824 (10th Cir. 1982).

28. Pouncy v. Prudential Ins. Co., 668 F.2d 795, 801 (5th Cir. 1982). "The disparate impact model requires proof of a causal connection between a challenged employment practice and the composition of the work force. Aptitude tests, height and weight requirements, and similar selection criteria all may be shown to affect one class of employees more harshly than another . . . ." Id. In Pouncy the plaintiff was unable to identify the selection device responsible for the lower representation of blacks in the employer's work force. Id. The statistics produced by the plaintiff could not be adequately connected to a challenged employment practice and the court refused to "require the employer to justify the legitimacy of any (or all) employment practices." Id. at 802.

29. E.g., Pouncy, 668 F.2d at 800 (the disparate impact model is not the "appropriate vehicle from which to launch a wide ranging attack on the cumulative effect of a company's employment practices"); But see Griffin v. Carlin, 755 F.2d 1516, 1523-24 (11th Cir. 1985)(disparate impact analysis is applicable to an overall employment process); Segar v. Smith, 738 F.2d 1249, 1271 (D.C. Cir. 1984)(criticizing the Pouncy court's position that complaints must pinpoint the specific cause of the discriminatory impact), cert. denied, 471 U.S. 1115 (1985).

30. Pouncy, 668 F.2d at 799-802. The court required "proof that a specific practice results in a discriminatory impact . . . in order to allocate fairly the parties' respective bur-
restriction on the application of the disparate impact analysis is inconsistent with the purpose of Title VII, considering the complexity of most employment practices.31

Once a prima facie case is established, the burden of proof is shifted to the defendant (employer), creating a presumption of unlawful discrimination.32 The employer now has the burden to show that legitimate, nondiscriminatory criteria were the basis of the selection process.33 This allocation is intended to focus the “inquiry into the elusive question of intentional discrimination.”34

The employer meets this burden by proving that the challenged criteria are manifestly related to the job.35 The plaintiff may respond by showing that alternate selection procedures were available, which were not discriminatory and which would also satisfy the legitimate interests of the employer.36

II. DEFENSES

Job relatedness,37 business necessity,38 or bona fide occupational qualification (BFOQ)39 are the only defenses available to an employer to uphold an employment practice which discriminates
dens of proof at trial.” Id. at 800.
33. Id. at 255. The employer's burden is one of production, not persuasion. Id. at 254. The explanation need only raise “a genuine issue of fact” as to whether the employer discriminated. Id. The burden of persuasion remains with the plaintiff throughout. Id. at 256.
34. Id. at 255 n.8.
Not withstanding any other provision of this subchapter, (1) it shall not be an unlawful employment practice for an employer to hire and employ employees, . . . on the basis of his religion, sex, or national origin in those certain instances where religion, sex, or national origin is a bona fide occupational qualification reasonably necessary to the normal operation of that particular business or enterprise . . . .
against a protected class.\textsuperscript{40}

A bona fide occupational qualification is a statutory exception to Title VII, but is limited to those instances which are "reasonably necessary to the normal operation of that particular business . . . ."\textsuperscript{41} This exception is further curtailed by the Equal Employment Opportunity Commission's (EEOC)\textsuperscript{42} guidelines on sex discrimination.\textsuperscript{43}

\textsuperscript{40} Traditionally courts would overlook a discriminatory employment practice if the employer would compensate by hiring a proportionate number of members of the protected class. See Blumrosen, The Bottom Line Concept in Equal Employment Opportunity Law, 12 N.C. CENT. L.J. 1 (1980). This concept is premised on Title VII's policy of increasing employment opportunities. \textit{Id.} at 5. The bottom line concept "protects employers who improve employment opportunities from 'direct' discrimination claims by minorities and women." \textit{Id.} at 5-6. This concept of correcting at the bottom line for discriminatory employment practices was rejected by the Supreme Court in Connecticut v. Teal, 457 U.S. 440 (1982).

As noted by the \textit{Teal} Court, Title VII guarantees the individual equal opportunity based on job-related criteria, by eliminating barriers which have a discriminatory impact. \textit{Id.} at 449. Equal opportunity is not facilitated by bottom line compensation. \textit{Id.} at 451. Congress never intended to allow employers to discriminate against some members of a protected group as long as other members benefit. \textit{Id.} at 455.


\textsuperscript{42} The EEOC was created, pursuant to 42 U.S.C. § 2000e-4(a), to provide leadership and coordination to the federal policy of establishing equal employment opportunity, pursuant to Exec. Order No. 12067, 43 Fed. Reg. 28967 (1978), \textit{reprinted in} 42 U.S.C. § 2000e app. at 32 (1982).

\textsuperscript{43} EEOC Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1604.2 (1986) provides in pertinent part:

Sex as a bona fide occupational qualification.

(a) The commission believes that the bona fide occupational qualification exception as to sex should be interpreted narrowly. Label—"Men's jobs" and "Women's jobs"—tend to deny employment opportunities unnecessarily to one sex or the other.

(1) The Commission will find that the following situations do not warrant the application of the bona fide occupational qualification exception:

(i) The refusal to hire a woman because of her sex based on assumptions of the comparative employment characteristics of women in general. For example, the assumption that the turnover rate among women is higher than among men.

(ii) The refusal to hire an individual based on stereotyped characterizations of the sexes. Such stereotypes include, for example, that men are less capable of assembling intricate equipment; that women are less capable of aggressive salesmanship. The principle of nondiscrimination requires that individuals be considered on the basis of individual capacities and not on the basis of any characteristics generally attributed to the group.

(iii) The refusal to hire an individual because of the preferences of coworkers, the employer, clients or customers except as covered specifically in paragraph (a)(2) of this section.

(2) Where it is necessary for the purpose of authenticity or genuineness, the Commission will consider sex to be a bona fide occupational qualification, e.g., an actor or actress.

29 C.F.R. § 1604.2(a) (1986).
which allow the BFOQ exception only to insure "authenticity or genuineness." The Supreme Court, relying on the restrictive language of the statute establishing the BFOQ, its legislative history, and the interpretation of the EEOC, has upheld this extremely narrow application of the BFOQ defense.

The BFOQ exception, as interpreted by the EEOC, requires a reasonable determination that all, or substantially all, women would be unable to perform the job safely and efficiently because strenuous manual labor is required, or where the same sex is required to accommodate the privacy of customers or clients. Sex discrimination can only be permitted if the business actually depends on the hiring of an exclusive sex. This necessity requirement is also the basis for the "business necessity" defense.

A "business necessity" may also justify an employment practice which has a disparate impact on a protected class. The existence of


45. Dothard v. Rawlinson, 433 U.S. 321, 334 (1977). Women working as prison guards in an institution where 20% of the male prisoners are sex offenders who are not segregated from the general inmate population would be subject to an increased risk of sexual assault. Id. at 335. This fact convinced the Court to accept the BFOQ defense offered by the state. Id. at 336-37. The Court did, however, reject the use of height and weight requirements and warned that expansion of this decision "beyond its narrow factual basis would erect a serious roadblock to economic equality for women." Id. at 347 (Marshall, J., concurring in part and dissenting in part).

46. Weeks v. Southern Bell Tel. & Tel. Co., 408 F.2d 228, 235 (5th Cir. 1969).

47. B. Schlei & P. Grossman, supra note 44, at 290. The authors cited City of Philadelphia v. Pennsylvania Human Relations Comm'n, 7 Pa. Commw. 500, 300 A.2d 97 (1973), as an example of a same sex BFOQ. In that case, youth counselors who were employed to conduct body searches and supervise children's showers were required to be of the same sex as the children. Id. at 291.


49. Griggs v. Duke Power Co., 401 U.S. 424, 431 (1971) (high school diploma requirement and intelligence test were not sufficiently related to the overall quality of the work force). In Griggs, the Supreme Court, while expanding the scope of potential liability under Title VII by adopting the "disparate impact theory," limited this approach by including the "business necessity" defense. Id. at 431. The Court in Griggs, however, failed to discuss the scope of the business necessity defense. Id. In New York City Transit Auth. v. Beazer, 440 U.S. 568 (1979), the Court construed the defense broadly. Id. at 594. Lower courts, however, have construed the defense narrowly. See Blake v. City of Los Angeles, 595 F.2d 1367, 1377 (9th Cir. 1979) (defining the defense as "very narrow"), cert. denied, 446 U.S. 928 (1980); Spurlock v. United Airlines, 475 F.2d 216, 219 (10th Cir. 1972) (relying on the uniquely high risks of air travel to establish the defense); United States v. Bethlehem Steel Corp., 446 F.2d 652, 662 (2d Cir. 1971) (adopting a "balancing approach"); see also infra notes 52-57 and accompanying text.
an "overriding legitimate business purpose," which is essential to the safe and efficient operation of the business, is the basic premise of this defense. The business necessity defense has three elements: (1) the business necessity must override the social harm which results from discrimination; (2) the employment practice must fulfill the necessary purpose it is alleged to serve; and (3) there must be no alternative practice which would better fulfill this need.

The amount of skill required by the prospective employment and the degree of human and economic risks were considered the decisive factors in the business necessity analysis. This analysis was subsequently restricted to those cases where "public safety is demonstrably jeopardized" and where the "selection criteria or employment condition is manifestly related to the protection of human safety." The risk of economic loss was no longer considered a major component of the business necessity test; rather safety became the determining factor.

This defense compares the benefits of the employment practice to the employer and society with the prohibitive cost of employment discrimination. Compelling benefits must be derived from a challenged employment practice in order to overcome society's desire to eliminate employment discrimination.

52. Robinson v. Lorillard Corp., 444 F.2d 791, 799 n.8 (4th Cir.), cert. dismissed, 404 U.S. 1006 (1971). If an alternate practice is available, then the one in question cannot be essential. Id. n.7. See Albemarle Paper Co. v. Moody, 422 U.S. 405, 425 (1975)(a showing of a reasonable alternative would be evidence that the selection procedure or policy was "merely a 'pretext' for discrimination").
55. Id.
57. See Spurlock v. United Airlines, 475 F.2d 216, 219 (10th Cir. 1972). Requirement that applicants for flight positions possess college degree and a minimum amount of experience are legitimate standards despite their discriminatory impact. The fact that airline pilots fly aircraft worth $20 million and transport as many as 300 passengers are important public interests which override the threat of discrimination. New York City Transit Auth. v. Beazer, 440 U.S. 568, 587 n.31 (1979) (Methadone users were uniformly denied employment as transit workers, the potential risk to public safety being too compelling); Burwell v. Eastern Air Lines,
An employment practice which discriminates against a protected class may also be justified if it can be shown to be job-related. The job-related defense is usually reserved for cases involving employment tests and requires the identified tasks of the job, which are the basis for the test, to be essential for actual job performance. The policy of this defense is to allow testing, but only if the test can adequately measure job performance. Tests "must measure the person for the job and not the person in the abstract." The relationship between the test and job performance must be validated, which requires the "minimum skill, ability, or potential necessary for the position at issue" to be ascertained. "[W]hether the qualifying tests are appropriate for the selection of qualified applicants for the job in question" must also be determined.

III. Test Validation

Test validation is the primary component of the job-related defense. The EEOC, which is the enforcement agency under the Civil

633 F.2d 361, 371-72 (4th Cir. 1980)(employer's mandatory pregnancy leave policy for flight attendants between 13th and 28th weeks of pregnancy, as well as mandatory leave from the beginning of the 28th week, was a legitimate business necessity to effect passenger safety), cert. denied, 450 U.S. 965 (1981); Note, Fair Employment Practices: The Concept of Business Necessity, 3 MEM. ST. U.L. REV. 76 (1972); Comment, The Business Necessity Defense to Disparate—Impact Liability Under Title VII, 46 U. Chi. L. REV. 911 (1979)(a detailed discussion of the restrictive application of the business necessity defense). But see Blake v. City of Los Angeles, 595 F.2d 1367, 1376-77 (9th Cir. 1979)(sex-segregated job classifications were not necessary to the safe and efficient operation of the Los Angeles Police Department; less discriminatory alternatives having been available thereby invalidating the claim of business necessity), cert. denied, 106 S. Ct. 1516 (1986); Gillespie v. State of Wisconsin, 771 F.2d 1035, 1040 (7th Cir. 1985), cert. denied, 106 S. Ct. 854 (1986); Craig v. County of Los Angeles, 626 F.2d 639, 662 (9th Cir. 1980), cert. denied, 450 U.S. 919 (1981); Allen v. City of Mobile, 466 F.2d 122 (5th Cir. 1972), cert. denied, 412 U.S. 909 (1973).


62. Id.

63. See infra notes 66-84 and accompanying text. Validation is a term of art designating the technical process of determining the job-relatedness of a selection procedure.


65. Id.

66. A validation study determines whether test results have a significant relationship to
Rights Act, has established guidelines to validate tests. These guidelines are based on professional standards of test validation established by the American Psychological Association (APA). They constitute "[t]he administrative interpretation of the Act by the enforcing agency," and therefore are indispensible to an employment test analysis. The EEOC guidelines are not, however, regulations promulgated pursuant to rulemaking procedures, and thus are not legally binding. Despite recent criticism, compliance with the EEOC guidelines is the most effective method to show that an employment test is job-related.

The EEOC guidelines list three methods of test validation: criterion-related, content, and construct. A criterion-related validity study requires "empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance." A content validity study requires "data showing that the content of the selection procedure is representative of important aspects of performance on the job for which the candidates are to be evaluated." A construct validity study requires "data showing that the procedure measures the degree to which candidates have identifiable characteristics which have been determined to be important in successful performance in the job for which the candidates are to be evaluated."
All three of these techniques must be consistent with recognized professional standards for evaluating tests, such as those published by the APA. All three validation techniques require specific documentation of the validity study, and the research must be conducted in an environment conducive to accuracy, which includes scoring of the tests under standardized conditions. The guidelines also specify that orientation, methods of selection, and cut-off scores can invalidate a test. Technical standards for each study are included in the EEOC guidelines.

Compliance with validation standards may not exonerate the employer, because a validated test may be considered discriminatory when alternate testing procedures exist. Validation requirements may also be less stringent in certain compelling situations. If a job requires a high degree of skill and "the economic and human risks involved in hiring an unqualified applicant are great, the employer bears a correspondingly lighter burden to show that his employment criteria are job-related." Few validation studies, however, ever satisfy the EEOC standard.

by identifying criteria that indicate successful job performance and then correlating test scores and the criteria so identified; "construct" validity (demonstrated by examinations structured to measure the degree to which job applicants have identifiable characteristics that have been determined to be important in successful job performance); and "content" validity (demonstrated by tests whose content closely approximates tasks to be performed on the job by the applicant).

77. 29 C.F.R. § 1607.5(C) (1986).
78. 29 C.F.R. § 1607.5(D)-(E) (1986).
79. 29 C.F.R. § 1607.5(F)-(H) (1986). Employers should avoid selection "on the basis of . . . knowledges, skills, or abilities . . . learned in brief orientation period . . . ." 29 C.F.R. § 1607.5(F) (1986). The selection device (pass/fail, rank order), which has an adverse impact, must be supported by sufficient evidence of validity to be used by the employer. 29 C.F.R. § 1607.5 (G) (1986). See infra notes 104-105 and accompanying text (Berkman use of Rank ordering). Furthermore, the use of cut-off scores must be closely examined and "the degree of adverse impact should be considered." 29 C.F.R. § 1607.5 (H) (1986).
80. C.F.R. § 1607.14 (D) (1986). Construct validity is a more complete strategy than either criterion-related or content validity. However, construct validation is a new and undeveloped procedure, and at present there is a substantial lack of literature related to employment testing.
82. Spurlock v. United Airlines, 475 F.2d 216, 219 (10th Cir. 1972)(position of airline pilot requires such a sliding scale because of the high risk involved).
83. Id.
84. Clady v. County of Los Angeles, 770 F.2d 1421, 1430 (9th Cir. 1985)(validation did not satisfy EEOC standard, but demonstrated legally sufficient correlation to success on the job to justify use by the county despite the disparate impact).
IV. PHYSICAL TESTS

Physical tests (or agility tests) possess unique validation problems which are compounded by social misconceptions regarding physical characteristics. A physical test that has a disparate impact on women cannot rely on the inherent differences between the male and female anatomy to validate its use, but rather the test must be shown to be job-related.86 Minimum height and weight requirements have been routinely invalidated by courts86 because they invariably discriminate against women and are not rationally based on job performance.87 The actual composition of the test is irrelevant. Courts focus on the disparate impact and the job-relatedness of the test.88

In 1978, New York City administered examination number 3040 for the purpose of selecting candidates for the entry level position of firefighters. The physical portion of exam 3040 was completed by 16,925 men and 79 women. Forty-six percent (46%) of the men passed, while none of the women did.88 As a result one of the female applicants instituted a lawsuit pursuant to Title VII. In Berkman v. City of New York,90 Brenda Berkman asserted that these results clearly established a prima facie case of discrimination, and

85. Blake v. City of Los Angeles, 595 F.2d 1367, 1375 (9th Cir. 1979).
86. See, e.g., Costa v. Markey, 706 F.2d 1, 6 (1st Cir. 1982) (height requirement, which was not shown to be job-related, discriminated against women), cert. denied, 464 U.S. 1017 (1983); United States v. Commonwealth of Virginia, 620 F.2d 1018, 1021 n.3, 1024 (4th Cir.) (height and weight requirements which eliminated 98% of all women, without evidence demonstrating a legitimate need for the requirements, contravene Title VII, and were voluntarily dropped by the state), cert. denied, 449 U.S. 1021 (1980); Blake v. City of Los Angeles, 595 F.2d 1367, 1374 (9th Cir. 1979) (height requirement that excluded 95% of all women and was not significantly related to police work and violates Title VII); Smith v. Troyan, 520 F.2d 492, 497 (6th Cir. 1975) (weight requirement discriminated against women, and was not rationally related to police work), cert. denied, 426 U.S. 934 (1976); Vanguard Justice Soc. v. Hughes, 471 F. Supp. 670 (D. Md. 1979) (height requirement excluded 81.8% of all female applicants in the age group of 18 to 34 years and defendants had not established the requirement to be job-related); Officers for Justice v. Civil Serv. Comm'n of San Francisco, 395 F. Supp. 378 (N.D. Cal. 1975) (height requirement had an adverse impact on women and had not been demonstrated as job-related). But see Jarrell v. Eastern Air Lines, 430 F. Supp. 884 (E.D. Va. 1977) (weight was subject to individual control and could be restricted by the employer), aff'd, 577 F.2d 869 (4th Cir. 1978).
88. Harless v. Duck, 619 F.2d 611, 614-16 (6th Cir.), cert. denied, 449 U.S. 872 (1980) (basic physical test, graded pass/fail, which was comprised of four parts and only three had to be completed in order to pass, was invalidated based on the disparate impact on female applicants and the inconclusive validation study).
90. 536 F. Supp. 177 (E.D.N.Y. 1982), aff'd, 705 F.2d 584 (2d Cir. 1983).
the burden shifted to the city of New York to show "that Exam 3040 was job-related, that is, that the test was valid because it accurately selected those applicants who will make better firefighters."\textsuperscript{93}

The ultimate issue in that case was whether a competitive physical exam can be sufficiently validated to justify such adverse results.\textsuperscript{92} Unfortunately, the design, administration, and validation technique of Exam 3040 were flawed due to the incompetent performance of representatives of the fire department and the city's Department of Personnel, which precluded the district court from fully examining the validation technique and thereby prevented the development of a workable rule.\textsuperscript{93}

The city had contracted with the American Institutes for Research (AIR) to develop physical test validation procedures.\textsuperscript{94} The contract was authorized under Intergovernmental Personnel Act Grant Number 72-14\textsuperscript{95} and was paid for by the U.S. Civil Service Commission.\textsuperscript{96} The city of New York, however, discarded significant portions of this federally funded expert advice, thereby undermining the cogency of the study.\textsuperscript{97} AIR considered criterion validation to be the appropriate method to validate the physical exams, but the city abandoned this strategy because of a fiscal crisis and relied on content validation.\textsuperscript{98} The city rejected tests for flexibility and equilibrium, citing their need for "firefighters, not for ballet dancers,"\textsuperscript{99} and substituted tests which compromised the comprehensive structure of the battery of tests suggested by AIR. The court, however, considered the test the city developed unreliable.\textsuperscript{100} The city's ill-advised decisions prevented the court from applying a reduced standard of job-relatedness, based on the intrinsic dangers of firefighting and its dedication to public safety.\textsuperscript{101}

\textsuperscript{91} Id. at 205-06.
\textsuperscript{92} Id. at 206-08.
\textsuperscript{93} Id. at 208.
\textsuperscript{94} Id. at 184.
\textsuperscript{96} 536 F. Supp. at 183.
\textsuperscript{97} Id. at 208.
\textsuperscript{98} Id. at 206-07.
\textsuperscript{99} Id. at 196.
\textsuperscript{100} Id. at 208-10. AIR contributed to the problems confronting the City's attempt to validate test 3040 by referring to tests as "face valid," an APA term of art, which is no substitute for validation based on job performance. Id. at 207.
\textsuperscript{101} See, e.g., Spurlock v. United Airlines, 475 F.2d 216 (10th Cir. 1972); Berkman, 536 F. Supp. at 209 n.20.
The district court questioned the use of a competitive test, which is intended to evaluate strength and speed, as an appropriate method to select firefighter candidates.102 This argument was supported by testimony which diminished the importance of speed and considered endurance to be the most reliable indicator of future job performance because performing at maximum speed may be hazardous in an actual fire.103 Safe technique and individual endurance, however, are characteristics which are acquired and are constant for any group of candidates. The city therefore had a legitimate interest in selecting candidates who have the greatest physical potential.

The district court criticized every aspect of Exam 3040, finding design and administrative flaws with each part of the exam.104 The futile attempt at validation and the use of rank order grading105 were

102. Berkman, 536 F. Supp. at 212. But see Vulcan Soc’y of the New York City Fire Dep’t v. Civil Serv. Comm’n of New York, 360 F. Supp. 1265 (S.D.N.Y.), aff’d, 490 F.2d 387 (2d Cir. 1973). In Vulcan Society, an earlier firefighter exam case, the written portion of the firefighter test was invalidated since it discriminated against blacks and Hispanics. The court considered a competitive physical to be a proficient means to select firefighter candidates, stating “[o]ne hardly needs an expert to be aware of the importance of physical capacity in performing the duties of a fireman.” Id. at 1276. The court went on to state that a competitive physical “might well have had a significantly favorable impact upon the relative performance of minorities . . . .” Id. at 1277. This decision and the Berkman dissatisfaction with competitive exams appear to be irreconcilable, and present New York City with an untenable situation of trying to resolve these two inconsistent decisions.


104. Id. at 200-04. The physical portion of Exam 3040 consisted of seven parts, all of which were criticized by the district court. First, the “dummy carry” required candidates to lift a cylindrical dummy, “without handles or other articulation” which presented a problem to those unfamiliar with the proper technique (only 4 of 80 women succeeded in lifting the dummy). Id. at 201. Second, the hand grip employed a device used to test hand strength which was adjustable to accommodate hand size, but this feature was not explained to candidates which handicapped those with small hands. Id. Third, free style broad jump used a mat which plaintiff claimed slipped. This and other flaws in the administration of the jump convinced the court to question the degree of precision used in the scoring. Id. at 201-02. Fourth, the flexed arm hang was used by the city despite AIR’s recommendation against it, since this test was considered by AIR to be unreliable. Id. at 202. Fifth, the agility test included an eight foot wall which deprived candidates of demonstrating abilities on the remainder of the test, since inability to negotiate the wall “automatically dropped one’s score from a possible 250 to 59.” Id. at 203. Sixth, the ledge walk received criticism similar to that of the broad jump, where flaws in administration, including lack of instruction on the wearing of the breathing apparatus, brought into question the degree of precision in the scoring. Id. at 203-04. Seventh, the mile run which was intended to measure stamina or aerobic capacity, was not considered by the court to be an accurate test for that purpose, since pacing was required. The court considered a test requiring a candidate to run or walk for twelve minutes and then to measure the distance to be more accurate. Id. at 204.

105. All applicants who passed both the written and physical portions of the exam were divided into 300 ranked groups. All of the candidates in the preceding group would have to be exhausted before anyone in the next group would be eligible eligible. Id at 211.
also severely censured by the district court.\textsuperscript{108} As a result, the city was ordered to prepare a new, validated selection procedure.\textsuperscript{107}

The new test, Exam 1162, was validated according to EEOC guidelines and APA standards to guarantee its neutrality.\textsuperscript{108} Both the EEOC guidelines and the APA standards are methods of validating an employment practice which has a disparate impact.\textsuperscript{109} In addition, Berkman, the plaintiff in both proceedings, her experts, and attorneys were invited to contribute to the preparation and validation of the new test, as a means to "prevent discriminatory impulses from realizing themselves in any blatant way in the process of test preparation."\textsuperscript{110} Exam 1162 was ruled to be job related on the basis of both a facial validation study, which is criterion-related,\textsuperscript{111} and a content validation study.\textsuperscript{112} Despite this arduous effort, the results of

\textsuperscript{106} Id. The EEOC is also dissatisfied with strict rank-order grading. EEOC Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607.14(C)(9) (1986).

\textsuperscript{107} Berkman v. City of New York, 705 F.2d 584, 588 (2d Cir. 1983). The city was ordered to hire up to 45 women, but eventually 38 women were hired. Id. at 592. No long range numerical requirements, however, were included in the decision. Id. at 597.


\textsuperscript{111} Id. at 595-96. The district court labelled one of the validation studies as a face valid study. Id. at 596. The appellants (the city and the Uniform Firefighters Association) contended this study is the more recognized criterion-related standard. EEOC Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607.5(B) (1986) (the standard employed is asserted to be in strict compliance with the EEOC Guidelines); Municipal Appellant's Brief at 16-17, Berkman v. City of New York, 812 F.2d 52 (2d Cir. 1987)(No. 86-7157); Brief for Defendant-Intervenor-Appellant Uniformed Firefighters Association at 25-27, Berkman v. City of New York, 812 F.2d 52 (2d Cir. 1987)(No. 86-7157). The study, conducted by Dr. Richard Reilly, demonstrated job performance with significant reliability. Municipal Appellant's Brief at 16. Exam 1162 predicted job performance at .01 levels of significance, which satisfied the EEOC Guideline requirement of a .05 level. Brief for Uniform Firefighters Association at 25-26. A .01 level means there is less than one chance in a hundred that the statistical relationship between the test and job performance occurred by chance. Id. Also, to guarantee fairness in the administration of the exam a fairness study was conducted, which again had positive results. Id. at 27.

\textsuperscript{112} Berkman v. City of New York, 812 F.2d 52, 56, 62 (2d Cir. 1987). See Gillespie v. State of Wisconsin, 771 F.2d 1035 (7th Cir. 1985), cert. denied, 106 S. Ct. 854 (1986). Content validation is the appropriate procedure to validate a test which measures "concrete characteristics in a form that simulated actual work behavior . . . ." Id. at 1043. This content validation adequately justifies the use of the challenged employment test despite the disparate
the new test, Exam 1162, had a disparate impact on women. Tests that employ strength as a criteria often fail to satisfy the validation requirements, but this trend was reversed by Exam 1162. This exam effectively tested upper body strength, which was shown to be a valid firefighting requirement. The district court, however, did not accept the emphasis placed on speed in the scoring model, noting its absence from the job analysis. This deficiency, combined with the dissatisfaction with the banded scoring technique employed by the city and the fact that the written portion of the exam failed

impact on a protected group.

113. Berkman v. City of New York, 812 F.2d 52, 55 (2d Cir. 1987) (21,236 men and only 77 women passed the physical portion of the test).

114. See Dothard v. Rawlinson, 433 U.S. 321, 331 (1977) (strength was not effectively related to job performance); Blake v. City of Los Angeles, 595 F.2d 1367, 1381-82 (9th Cir. 1979) (upper body strength requirement of police candidates was not shown to be job-related); Berkman v. City of New York, 536 F. Supp. 177, 212 (E.D.N.Y. 1982) (tests of maximum strength are not determinative to effective job performance as a firefighter), aff'd, 705 F.2d 584 (2d Cir. 1983).


117. Id.

118. Id. at 600-01. The city's description of the scoring model is as follows: a panel consisting of Fire Department supervisors tested 92 incumbent firefighters with Exam 1162. Profiles describing acceptable, unacceptable; and exemplary levels of performance were developed and provided to assist in the evaluation. Municipal Appellants' Brief at 9-11, Berkman v. City of New York, 812 F.2d 52 (2d Cir. 1987)(No. 86-7157). A passing grade was established from this procedure. Based on the results of the panel members' evaluation of the performance of the incumbents, expert recommendations, and objective life saving criteria, the city developed a seven band scoring system based on thirty second intervals.

<table>
<thead>
<tr>
<th>Minutes Test Score</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4:00 and under 100</td>
<td>100</td>
</tr>
<tr>
<td>4:01 to 4:30 95</td>
<td></td>
</tr>
<tr>
<td>4:31 to 5:00 90</td>
<td></td>
</tr>
<tr>
<td>5:01 to 5:30 85</td>
<td></td>
</tr>
<tr>
<td>5:31 to 6:00 80</td>
<td></td>
</tr>
<tr>
<td>6:01 to 6:30 75</td>
<td></td>
</tr>
<tr>
<td>6:31 to 7:00 70</td>
<td></td>
</tr>
<tr>
<td>OVER 7:00 FAILS</td>
<td></td>
</tr>
</tbody>
</table>

Municipal Appellants' Brief at 11.

All candidates within a band would be randomly selected (to minimize the effect of minor scoring deviations) and all candidates in the preceding band must be exhausted before the next band is started. Municipal Appellants' Brief at 11; See Berkman, 536 F. Supp. at 201-04.

The physical test consisted of the following series of tasks:

- Part I - Engine Work Simulation, which included hose drag, hose carry up stairs and hose pull.
- Part II - Mandatory Walk.
- Part III - Ladder Work Simulation, which included scaling a 4.5 foot wall, ladder raise, ladder ascent and descent, ascending stairs with a 15 pound dumbbell weight,
to distinguish between the cognitive abilities of the candidates, persuaded the district court to effectuate changes in the scoring of Exam 1162.\textsuperscript{119}

In an effort to reduce the adverse impact on women, the district court ordered three changes in the scoring system.\textsuperscript{120} The district court noted that a disproportionate number of female candidates elected not to participate in the physical portion of Exam 1162 and attributed this trend to the "continued discrimination within the Fire Department . . . ."\textsuperscript{121} As a method of compensation, the city was ordered to employ a selection ratio which would facilitate the hiring of women over equally qualified males.\textsuperscript{122}

The district court also instructed the city to discard the written test results and use computer generated random scores in their place,\textsuperscript{123} because the "bunching of the written test scores" would detrimentally increase the importance of the physical portion of the exam.\textsuperscript{124} A more difficult written exam, which successfully distinguished the cognitive abilities of the candidates, however, would probably subject itself to a Title VII challenge requiring a validation of the job-relatedness of cognitive abilities.\textsuperscript{125} The district court also

\textsuperscript{119} Berkman v. City of New York, 812 F.2d 52, 57 (2d Cir. 1987).

\textsuperscript{120} Id.

\textsuperscript{121} Id. The district judge considered discrimination in the fire department to be evidenced by "the well-publicized efforts of the department to discharge the plaintiff and another female probationary firefighter in September 1983." Id. The district judge assumed that "in the absence of the deterrent effect of the attempted firings, the percentage of those taking the physical test after passing the written would have been the same for women as for men . . . ." Id. On this premise, the district judge calculated the selection ratio to compensate. Berkman v. City of New York, 626 F. Supp. 591, 601 (E.D.N.Y. 1985).


\textsuperscript{123} Berkman v. City of New York, 812 F.2d 52, 58 (2d Cir. 1987) (computer generated random scoring is an electronic method to randomly allocate new scores to candidates, which maintains the same level of fairness that is attributed to any lottery system).

\textsuperscript{124} Berkman v. City of New York, 812 F.2d 52, 56, 58 (2d Cir. 1987) The median score of the written test was between 96 and 97, and 9,788 candidates received a score of 98 or above. Berkman, 626 F. Supp. at 594. Since the written portion of Exam 1162 was not effective in distinguishing the cognitive abilities of the candidates, the physical portion became determinative. In order to rectify this inequitable situation the court suggested allocating to each candidate a random score to diminish the effect of the physical and to compensate for the discriminating effects of adverse publicity. Id. at 601. See also supra note 102.

\textsuperscript{125} See Vulcan Soc'y of Westchester County v. Fire Dep't of White Plains, 505 F. Supp. 955, 965-66 (S.D.N.Y. 1981)(invalidating the use of a high school diploma as a fire department employment requirement); Vulcan Soc'y of the New York City Fire Dep't v. Civil Serv. Comm'n, 360 F. Supp. 1265, 1275-76 (S.D.N.Y.)(invalidating written exam testing
criticized the banded scoring technique employed by the City, and suggested the use of three broader bands to reduce the adverse impact on women.\textsuperscript{128}

The United States Court of Appeals for the Second Circuit reversed the orders of the district court concerning scoring adjustments and allowed the city to implement the results of exam 1162.\textsuperscript{127} The court of appeals considered the replacement of the seven band scoring model with a three band system to be "fatally flawed," since this adjustment "neither enhances the validity of the physical test nor reduces the adverse effect on women applicants."\textsuperscript{128} The court of appeals considered the use of computer generated written scores to be an unwarranted remedy, which "unfairly deprives many male and female applicants of the enhanced opportunity they achieved by scoring comparatively better on the written test than other applicants."\textsuperscript{128} This remedy also burdens the fire department with applicants who demonstrated a lesser degree of cognitive ability.\textsuperscript{130} The court of appeals also rejected the compensation ratio since its effect would be extremely limited.\textsuperscript{131}

The validity of the physical portion of Exam 1162 was also contested in a cross appeal filed by Ms. Berkman,\textsuperscript{132} but the court of appeals upheld the validity of the test on the basis of content validation and criterion-related validation.\textsuperscript{133} This decision may establish a viable validation standard for competitive physical tests, which can be relied on by other municipalities.

\begin{itemize}
\item knowledge of New York City government and current events), \textit{aff'd}, 490 F.2d 387 (2d Cir. 1973).
\item 126. \textit{Berkman v. City of New York}, 812 F.2d 52, 57 (2d Cir. 1987).
\item 127. \textit{Id. at} 62-63.
\item 128. \textit{Id. at} 60. Under either system, whether the three band suggested by the court or the seven advocated by the city, only two (2) women would be reached for appointment; and that women would be reached sooner under a seven band scoring system. Municipal Appellants' Brief at 20, \textit{Berkman v. City of New York}, 812 F.2d 52 (2d Cir. 1987)(No. 86-7157). Also, under the proposed three band scoring system, no women would be in the top band, the one which must be exhausted first. This band would consist of 6,790 men. The second band would be comprised of 13,836 men and 44 women. These figures are based on plaintiff's consultant, Dr. Raymond Mendel, as represented in the Municipal Appellants' Brief at 18.
\item 129. \textit{Berkman}, 812 F.2d at 61.
\item 130. \textit{Id.}
\item 131. \textit{Id. at} 62 (the compensation ratio would only effect two women, since this remedy would not increase scores, merely establish a preference for women with the same score, which effectively would be two).
\item 132. \textit{Id. at} 58-59.
\item 133. \textit{Id. at} 59-60.
\end{itemize}
V. THE COST OF COMPETITIVE TESTING

The cost to institute an employment practice which will be effective and withstand Title VII scrutiny is surprisingly high. Tests must be validated, but as illustrated in the 1162 situation, this can be an expensive process because of the continued court appearances, and an ultimately useless process considering the technical requirements of the EEOC guidelines. Furthermore, a test which has to meet such rigorous guidelines is expensive to administer. The cost to administer the physical portion of Exam 1162 alone was $750,000. In the event the test is invalidated, the costs increase dramatically. The high cost of litigation, the appeals, and the opposing attorney's fees, all must be considered. Furthermore, the typical remedy provided by a victory in a Title VII litigation is backpay, which can be an expensive sanction and has been incorporated into the statute for that very reason.

The final cost in this analysis is less obvious, but ultimately more important. As a reaction to this expensive process of validating merit tests, municipalities may rely on less accurate methods of ascertaining the most qualified candidates for civil service employment. Employment practices to determine candidates for municipal services may be reduced to quotas. These quotas may curtail the quality of emergency services provided by police and fire departments, thus creating a hazard to the community and a danger to fellow workers.

134. See supra notes 73-80 and accompanying text. B. SCHLEI & P. GROSSMAN, supra note 44, at 181 (concluding that besides being expensive, test validation according to the EEOC guidelines may be impossible).

135. Berkman, 812 F.2d at 55.

136. See Follett & Welch, Testing For Discrimination in Employment Practices, 46 LAW & CONTEMP. PROBS., Autumn 1983, at 171. EEOC budget was $142 million in 1981, and in the same year 13,750 civil rights cases were tried in federal courts. The number tried under Title VII is not ascertainable, but 41% of the civil rights cases filed were brought under Title VII. Id.

137. See Sprogis v. United Air Lines, 517 F.2d 387, 392 (7th Cir. 1975)(attorney's fees are discretionary, but may be four times the amount of the damage award if the case is of substantial precedential value); Albemarle Paper Co. v. Moody, 422 U.S. 405, 415 (1975)(the strong public interest to eliminate discriminatory employment practices supports the practice of awarding attorney's fees).

138. See supra notes 16-18 and accompanying text.


VI. ALTERNATIVES

Assessment centers, which confront applicants with real life situations and assess their responses, are considered a viable alternative to the traditional employment test.\textsuperscript{141} An employment practice which included the use of an assessment center has been validated as criterion-related.\textsuperscript{142} The major drawback to this employment procedure is the prohibitive cost\textsuperscript{143} which has restricted its use until the later stages in the selection process when the applicant pool has been substantially reduced.\textsuperscript{144} The concern over the increased cost of assessment centers, however, must be compared to the aggregate cost of a traditional test, such as those involved in Berkman.\textsuperscript{145}

While the use of assessment centers is becoming an alternative to standardized tests,\textsuperscript{146} the actual test situation may not satisfy the validation criteria. The validity of a particular assessment center is dependent upon the duration of a candidate’s examination and the quality of the assessors, which is contingent on the extent of the assessors’ training and experience.\textsuperscript{147} Although the assessment center may more accurately predict job performance, its use as an employment practice would still retain the need for an approved method to rank candidates.\textsuperscript{148}

The city of Seattle took an alternate approach with respect to

\begin{itemize}
\item \textsuperscript{142} Rivera v. City of Wichita Falls, 665 F.2d 531 (5th Cir. 1982).
\item \textsuperscript{143} Note, \textit{Challenges to Employment Testing Under Title VII: Creating “Built in Headwinds” for the Civil Service Employer}, 12 \textit{FORDHAM URB. LI.}, 749, 771 n.114 (1984). The average assessment cost $540.25 per candidate, while the estimated cost of a pencil and paper test range from $10.00 to $68.82. \textit{Id}.
\item \textsuperscript{144} Guardians Ass’n of New York v. Civil Serv. Comm’n, 630 F.2d 79, 97 (2d Cir. 1980) (the use of assessment centers is too expensive for large numbers of candidates, but is useful after a written exam or background check has significantly reduced the number of applicants), \textit{cert. denied}, 452 U.S. 940 (1981).
\item \textsuperscript{145} \textit{See supra} notes 134-40 and accompanying text.
\item \textsuperscript{147} Firefighters Institute v. City of St. Louis, 616 F.2d 350, 360-62 (8th Cir. 1980).
\item \textsuperscript{148} \textit{See} Berkman v. City of New York, 536 F. Supp. 177, 211 (E.D.N.Y. 1982) (rank order selections can only be justified by a close correlation between the test and future job performance).
\end{itemize}
its fire department physical. Refusing to compromise the department's standards, Seattle groomed women to pass the traditional rigorous physical, but the price was high. In fiscal year 1982, Seattle spent $10,000 on each female firefighter candidate. Such an extensive training program, however, may be too expensive for a large municipality.

A final alternative is to utilize a pass/fail employment test, and include a random selection process to differentiate among the large numbers of candidates who will pass an unchallenged test. This method does not adhere to the articulated commitment to merit testing and may increase the risk of hiring unqualified candidates. These disadvantages could be circumvented by developing stringent training programs, which may ultimately be more cost effective than the present testing procedures.

CONCLUSION

Employment tests attempt to predict future job performance, a process that can never be completely accurate. A reasonable relationship between the job and the test must suffice. A strict application of the EEOC guidelines may "leave the employer little choice, save an impossibly expensive and complex validation study, but to engage in a subjective quota system of employment selection," which does not effectuate the purpose of Title VII, and may compromise the quality of the personnel. Simple quota hiring is even more attractive to the municipal employer, who is restricted by limited resources. The municipal employer may also be confronted with a severe personnel shortage. While the constant effects of attrition and retirement erode the workforce, efforts to replace these employees are delayed by Title VII litigation.

The application of Title VII is holding municipal employers to

150. Id. at 102.
151. See Berkman v. City of New York, 812 F.2d 52, 55 (2d Cir. 1987) (31,421 applicants took the written portion of Exam 1162).
152. See supra note 124 and accompanying text (results of Exam 1162, where almost everyone passed the written portion (98.58%)).
154. Id.
156. See Vulcan Soc'y of the New York City Fire Dep't v. Civil Serv. Comm'n, 360 F. Supp. 1265, 1278 (S.D.N.Y.), aff'd, 490 F.2d 387 (2d Cir. 1973).
an unrealistic standard. The EEOC guidelines need to be altered, or at least clarified, to enable a municipality to develop a job-related physical examination which will survive Title VII scrutiny, without carrying the exhausting burden which confronted the city of New York. The eight year struggle which characterized the New York City firefighter test must not be the standard of test validation. If these changes cannot be accomplished and the situation remains the same, municipalities should re-evaluate this senseless legal battle, since the entire process is far too expensive. The development and administration of the test, the litigation cost involved in defending its validity, and finally discarding the test and replacing it with a quota system, with the plaintiff seeking backpay and attorney’s fees, is far too high a price to pay for tests which many claim are unrelated to future job performance.

James M. Conway