U.C. DAVIS LAW REVIEW

University of California Davis

VOLUME 19 SUMMER 1986 NUMBER 4

ARTICLES

Rank-Order Physical Abilities Selection Devices for Traditionally Male Occupations as Gender-Based Employment Discrimination

*Ruth Colker**

Historically, women, as well as racial minorities, have been excluded from police and firefighter positions. Only recently, since the 1972 amendments to Title VII of the Civil Rights Act of 1964 extended coverage to state and local governments, have many police and fire departments begun to permit women to apply for employment. However, these departments also have instituted rank-order physical abilities selection devices that have effectively perpetuated the exclusion of women from police and firefighter jobs. This Article critically examines these selection devices. It notes that many police and fire departments assume the only physical demands of these traditional male activities involve the male-valued attributes of speed and strength. Yet society has denied women the

*Associate Professor of Law, Tulane University. A.B. 1978, Harvard-Radcliffe College; J.D. 1981, Harvard Law School. The author wishes to thank Mary Whisner, George Strickler, Joel Friedman, and John Dzienkowski for comments on earlier drafts, and Susan Dellar Ross for the valuable encouragement she provided the author when the author litigated this area while employed at the United States Department of Justice. The conclusions, of course, are the author's.

same opportunities to develop physical skills that men have enjoyed. This Article argues that valid exams must test for physical abilities actually related to job performance. Furthermore, an employer using a physical abilities selection exam on a rank-order basis must demonstrate that greater physical capabilities equal better job performance. Finally, this Article considers alternate selection devices, related to job performance, to bring women into the police and firefighter fields in greater numbers.

INTRODUCTION

Hiring police officers or firefighters to include blacks, Hispanics, Asians, and women poses a challenge for many municipalities. Historically, both women and minorities have been excluded from these professions.[1] After such absolute prohibitions were successfully challenged in 1973,[2] many governmental employers imposed rank-order written examinations, minimum height and weight requirements, and rank-order physical abilities examinations.[3] These selection devices have also excluded women and minorities: written examinations have disparately excluded blacks and Hispanics;[4] minimum height and weight requirements have disparately excluded Hispanics, Asians, and women;[5] and physical abilities examinations have disparately excluded women.[6]

The discriminatory aspects of rank-order written examinations and minimum height and weight requirements have been successfully challenged.[7] However, the use of rank-order physical abilities tests remains an employment obstacle for women. Although some police departments no longer use physical abilities tests as the result of court orders or consent decrees,[8] many of these departments may reinstitute these selection devices due to the United States Department of Justice's recent challenges to the validity of the orders and decrees.[9] Moreover, the widespread use of rank-order physical abilities examinations for

---

[1] *See infra* text accompanying notes 61-66.

[2] *See infra* note 61. These cases were generally brought under Title VII of the Civil Rights Act of 1964, 42 U.S.C. §§ 2000e-2000e-17 (1976 & Supp. 1979), which Congress amended in 1972 to cover state and local governments. For further discussion, see *infra* text accompanying note 217.

[3] *See infra* text accompanying notes 67-68. *Cf.* Johnson v. Mayor & City Council, 53 U.S.L.W. 4754, 4757 (June 17, 1985) (age requirement for firefighters).

[4] *See infra* note 95.

[5] *See infra* note 67.

[6] *See infra* note 72.

[7] *See infra* notes 67 & 95.

[8] *See infra* text accompanying notes 215-16.

[9] *See infra* text accompanying notes 11-19.

firefighter selection remains largely unchallenged despite their exclusionary impact on women's employment opportunities.[10]

A recent case, *United States v. District of Columbia*,[11] exemplifies the historical exclusion of women from firefighting. It also demonstrates the United States Justice Department's disregard of women's present employment opportunities in firefighting.[12] In *District of Columbia*, the United States alleged that the District of Columbia Fire Department discriminated against white men on the basis of race and sex in hiring and promotion through the use of a race- and sex-conscious affirmative action plan.[13] Ironically, the parties' stipulated facts

---

[10] *See infra* text accompanying note 135.

[11] United States v. District of Columbia, No. 85-0797 (D.D.C. filed Mar. 11, 1985). By focusing on the sex discrimination aspects of this case, the author does not mean to suggest that the race discrimination allegations are valid.

[12] This case received widespread attention because it was the first time the United States brought an employment discrimination case on behalf of white men. *See* cases cited *infra* note 95. *But see* United States v. Jefferson County, No. 75-P-0666-S (N.D. Ala. Feb. 27, 1984) (United States alleged that the city's failure to use a valid written examination on a rank-order basis discriminated against white applicants for certain firefighter positions).

By contrast, virtually no attention has been given to the sex discrimination aspects of this case. None of the parties to the case focused on the United States' sex discrimination allegation. For instance, the Statement of Stipulated Material Facts contains statistics that break down appointments by race, but does not contain statistics that break down appointments by sex. Hammon v. Barry, 606 F. Supp. 1082 (D.D.C. 1985). The Stipulated Facts also provide no statistics regarding how many women took the 1984 written examination, although they provide examination statistics on the basis of race. The court also ignored the allegation of sex discrimination. Although its opinion discussed at length the racial composition of the fire department, the court never discussed the sexual composition. *Hammon*, 606 F. Supp. at 1086-87.

Finally, the District of Columbia did not file a motion to dismiss the United States' allegation of sex discrimination despite the United States' failure to set forth a prima facie case of discrimination. *See infra* note 15 (defining prima facie case).

[13] United States' Complaint for Injunctive Relief, United States v. District of Columbia, No. 85-0797 (D.D.C. filed Mar. 11, 1985). This case began when four black applicants to the District of Columbia Fire Department filed suit in district court seeking enforcement of the city's affirmative action plan. *Hammon*, 606 F. Supp. 1082. On May 23, 1984, the court approved a consent decree in that case which required the District to submit an affirmative action plan (AAP) to the court for approval. On February 7, 1985, the District submitted to the court an AAP it had already put into effect. Eight white incumbent firefighters and their union then filed a complaint against the District alleging that the promotional provisions of the AAP were unlawful and unconstitutional. Byrne v. Coleman, No. 85-0782 (D.D.C. filed Mar. 8, 1985).

The United States, through the Attorney General, also filed a complaint against the District alleging that the hiring and promotion aspects of the AAP were unlawful or unconstitutional by requiring preferences based on race, color, or sex. United States v.

indicate that the District historically excluded women, not men, from firefighter positions; the stipulations also indicate that the affirmative action plan would not significantly improve women's opportunities for employment.[14] Rather than establish a prima facie case of discrimination against men on the basis of gender, the stipulated facts could have established a prima facie case of gender discrimination against women.[15]

District of Columbia represents far more than an inattention to women's employment opportunities. The case also demonstrates the growing popularity of rank-order physical abilities examinations that perpetuate the historic, absolute exclusion of women from firefighter positions.[16] Although the District of Columbia has never used a physical abilities examination to select firefighters, the United States suggested the District use such an examination on a rank-order basis as its

---

District of Columbia, No. 85-0797 (D.D.C. filed Mar. 11, 1985). In a March 14, 1985 order, the court consolidated the three cases. On April 1, 1985, the court found the promotional aspects of the AAP unlawful. It also found that the hiring aspects of the plan were legal and constitutional but that the District should on remand consider alternatives to the hiring plan. The United States has appealed the district court's decision.

[14] *See infra* note 15.

[15] In order to state a prima facie case of discrimination under Title VII, a plaintiff must demonstrate that the challenged employment practices had a disparate impact on the basis of a protected category (*i.e.*, race, sex, color, religion, or national origin). Griggs v. Duke Power Co., 401 U.S. 424, 431-32 (1971); *see also* Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607.4D (1985) [hereafter Uniform Guidelines]. According to the Uniform Guidelines, a plaintiff has stated a prima facie case of discrimination if she can demonstrate that she was less than four-fifths as likely to be hired as an equally qualified man.

The United States did not establish a prima facie case because the stipulated facts showed that men constituted 93% of the applicants who passed the District's examination and 95% of the individuals who would be selected under the District's AAP. Those statistics suggest that men fared *better* under the AAP than they would have through gender-neutral hiring under the District's examination. Moreover, women constituted seven percent of the individuals who passed the selection examination and only five percent of the individuals who would be hired under the AAP. *See* Statement of Stipulated Material Facts, Hammon v. Barry, No. 84-0903 (D.D.C. filed Mar. 14, 1985). A woman was therefore five-sevenths as likely to be hired as an equally qualified man. Under the four-fifths rule, it was *women*, not men, who could state a prima facie case of discrimination under the District's proposed affirmative action plan.

[16] *See infra* text accompanying notes 125-214; *see also* W. MCARDLE, F. KATCH & V. KATCH, EXERCISE PHYSIOLOGY 227-28, 288-91 (1981) (describing disparate impact that cardiac output and strength tests would have on women) [hereafter W. MCARDLE].

sole selection device.[17] The Justice Department proposed this hiring practice even though it admitted that the rank-order physical examination would "tend to adversely affect the employment opportunities available to women."[18]

The United States' justification for its proposal was that a rank-order physical performance test would be a "less restrictive alternative"[19] to the race-conscious hiring goals set forth in the District's affirmative action plan. Thus, the United States implicitly argued that the perpetuation of sex discrimination against women is an acceptable solution to the problem of race discrimination.

This Article argues that rank-order physical abilities examinations are not a lawful "alternative" to race-conscious hiring goals in traditionally male occupations.[20] Part I explores the stereotypes underlying the historical segregation of physical activity on the basis of gender in athletics and employment.[21] Part II discusses Title VII's validation requirements for physical abilities examinations that have a disparate impact on the basis of sex.[22] Part III applies these requirements to the

---

[17] United States' Memorandum Concerning Jurisdiction and Alternative Selection Procedures, Hammon v. Barry, No. 84-0903 (D.D.C. filed Apr. 25, 1985) [hereafter United States' Memorandum]. The district court was first to suggest that the District consider alternatives to its affirmative action plan, such as a physical abilities test. *Id.*, slip op. at 38-39 (Apr. 1, 1985). In response, the United States argued that a physical abilities test would be an acceptable alternative. United States' Memorandum, *supra*, at 7-8.

[18] United States' Memorandum, *supra* note 17, at 7 n.4. For further discussion of the adverse impact that such examinations have had upon women, see *infra* notes 125-214.

[19] According to the Uniform Guidelines, an employer should investigate "suitable alternative selection procedures and suitable alternative methods of using the selection procedure which have as little adverse impact as possible, to determine the appropriateness of using or validating them in accord with [the] guidelines." Uniform Guidelines, *supra* note 15, § 1607.3B.

[20] Black and Hispanic women are precluded from employment by both written examinations that have a disparate impact on the basis of race and physical performance examinations that have a disparate impact on the basis of sex. Hence, the United States' suggested alternative would benefit only black and Hispanic men at the expense of black and Hispanic women, as well as white women.

[21] *See infra* text accompanying notes 27-93. For further discussion of the effect of biologically based stereotypes on women's employment opportunities, see Bartlett, *Pregnancy and the Constitution: The Uniqueness Trap*, 62 CALIF. L. REV. 1532 (1974); Scales, *Towards a Feminist Jurisprudence*, 56 IND. L.J. 375 (1981); Segal, *Sexual Equality, The Equal Protection Clause and the ERA*, 33 BUFFALO L. REV. 85 (1984); Williams, *Equality's Riddle: Pregnancy and the Equal Treatment/Special Treatment Debate*, 13 N.Y.U. REV. L. & SOC. CHANGE 324 (1984-1985).

[22] *See infra* text accompanying notes 94-124.

stereotypes that underlie the use of rank-order physical abilities examinations for the selection of police officers and firefighters. It argues that perpetuation of women's exclusion from these occupations through the use of rank-order physical abilities examinations derives from gender-based stereotypes about men's and women's physical abilities and proper roles.[23] In its final part, the Article proposes an alternative selection device for the hiring of qualified firefighters or police officers that would neither have a disparate impact upon a protected group[24] nor require goals and quotas.[25]

The development of nondiscriminatory selection devices for police and firefighters is important; these public service occupations have a devastating history of discrimination based on race, national origin, and gender. Many police and fire departments have considered rank-order physical abilities tests the easiest and best solution to the historical problem of employment discrimination against black, Asian, and Hispanic men, because these examinations produce no disparate impact on the basis of race or national origin. Yet, that solution is unacceptable and often unlawful because it imposes a high price on women of all races and national origins.[26]

I. Socially Created Gender-Based Differences in Physical Opportunity

Police and fire departments use rank-order physical abilities tests because the departments assume higher test scores predict better job performance.[27] However, many physical abilities tests do not measure the attributes or skills specific jobs require.[28] As a consequence of socially created differences in men's and women's opportunities to develop

---

[23] *See infra* text accompanying notes 125-214.

[24] *See infra* text accompanying notes 215-32.

[25] By suggesting a selection device that would not require the use of goals and quotas, the author does not mean to suggest that goals and quotas are inappropriate or unlawful. The selection procedure proposed in this Article might result in nondiscrimination; however, it would not result in affirmative action on behalf of a group that previously has been excluded from employment. For example, if a fire department has never employed women, then it might be appropriate to use goals and quotas to achieve immediately a more than token representation of women in the fire department. After this fire department had achieved a substantial representation of women through affirmative action, then it might be able to turn to a nondiscriminatory selection device such as the one proposed in this Article.

[26] *See supra* note 20.

[27] *See infra* notes 186-214 and accompanying text.

[28] *See infra* notes 125-85 and accompanying text.

physical skills, men often rank higher in traditional physical abilities tests. Women historically have had little involvement in athletics and employment perceived as physically demanding, due to stereotypical assumptions about each sex's physical abilities.

A. Gender-Based Segregation in Athletics

Until the mid-1800's, American women often were valued for their strength and valor.[29] Nevertheless, it was important to society that middle-class women maintain their "femininity" while they performed "masculine" tasks. Writing in 1710, William Byrd described an ideal woman of those times:

> She is a very civil woman and shews nothing of ruggedness, or Immodesty in her carriage, yett she will carry a gunn in the woods and kill deer, turkey, &c., shoot down wild cattle, catch and tye hoggs, knock down beeves with an ax and perform the most manfull Exercises as well as most men in those parts.[30]

Strenuous physical activity became less acceptable for middle-class women as the United States became more industrialized. A "cult of domesticity" developed for middle-class women which maintained that their proper place was within the home, serenely domestic.[31] Society

---

[29] *See, e.g.*, J. SPRUILL, WOMEN'S LIFE & WORK IN THE SOUTHERN COLONIES (1972).

[30] *Id.* at 81 (quoting Byrd, *Boundary Line Proceedings*, 5 VA. MAG. 10 (1710)).

[31] *See generally* B. EHRENREICH & D. ENGLISH, COMPLAINTS AND DISORDERS: THE SEXUAL POLITICS OF SICKNESS (1973); Barker-Benfield, *Sexual Surgery in Late-Nineteenth Century America*, in SEIZING OUR BODIES: THE POLITICS OF WOMEN'S HEALTH 13 (C. Dreifus ed. 1978); Smith-Rosenberg & Rosenberg, *The Female Animal: Medical and Biological Views of Woman and Her Role in Nineteenth-Century America*, 60 J. AM. HIST. 335 (1973) [hereafter Smith-Rosenberg].

Samuel Jennings wrote in 1808 that:

> Men and women appear to best advantage each in their own proper station. Had it been my lot, to have taken one of those *manlike ladies*, whenever there happened to be company at my house, I should have made it my business, to brush the floor, rub the furniture, wash the tea equipage, scold the maids, talk about the kitchen and dairy, &c. and apologize as I proceeded, by giving intimation, that I had made an exchange of provinces with my good wife, by way of mutual accommodation. Such conduct would at least shew, how awkwardly a man appears in acting the part of a woman, and of course would lead a woman of common sense to conclude, that she could not appear to much better advantage, when engaged in the capacity of a man.

Jennings, *The Married Lady's Companion, or Poor Man's Friend*, *reprinted in* ROOT OF BITTERNESS: DOCUMENTS OF THE SOCIAL HISTORY OF AMERICAN WOMEN 116 (N. Cott ed. 1972).

encouraged middle-class women to avoid physical activity that might lead to "weak and degenerate offspring."[32] Working class and slave women who engaged in heavy physical labor were not considered feminine.[33]

While society encouraged women to be more feminine, it gave men the opposite message. In the 1860's athletics began to develop for men as a way to counteract the "feminizing" of male culture.[34] Women's athletics also began to develop, but the impetus for change came from quite different sources — an upper-class need for leisure activity, a middle-class development of sports at women's educational institutions, and a concern for healthier Caucasian children to "save" the race.[35] Although poor women continued to engage in physically demanding work, society did not give them the opportunity to engage in emerging middle-class athletics.[36]

The sports developed for women reflected the prevailing view that middle-class women could not participate in arduous physical activity. Basketball rules minimized exertion and physical contact, tennis tournaments were limited to three sets, and women's athletic clothing was too restrictive to permit much exertion.[37] Moreover, society promoted only sports considered appropriately feminine; it uniformly disapproved "masculine" sports such as baseball, boxing, football, and cross country

---

[32] Smith-Rosenberg, *supra* note 31, at 335.

[33] *See generally* Twin, *Introduction*, in OUT OF THE BLEACHERS: WRITINGS OF WOMEN AND SPORT xviii-xix (S. Twin ed. 1979) [hereafter OUT OF THE BLEACHERS]. Sojourner Truth, a freed slave, spoke strongly in 1851 about the contrast between social stereotypes about women's proper place and the reality of a black woman's physical accomplishments:

> The man over there says women need to be helped into carriages and lifted over ditches, and to have the best place everywhere. Nobody ever helps me into carriages or over puddles, or gives me the best place — and ain't I a woman? Look at my arm! I have ploughed and planted and gathered into barns, and no man could head me — and ain't I a woman? I could work as much and eat as much as a man — when I could get it — and bear the lash as well! And ain't I a woman?

Id. at xix (citing E. FLEXNER, CENTURY OF STRUGGLE: THE WOMEN'S RIGHTS MOVEMENT IN THE UNITED STATES 90-91 (1970)). For an excellent discussion of the historical treatment of black women, see J. JONES, LABOR OF LOVE, LABOR OF SORROW: BLACK WOMEN, WORK, AND THE FAMILY FROM SLAVERY TO THE PRESENT (1985).

[34] Twin, *supra* note 33, at xxi-xxii.

[35] *Id.* at xxiv.

[36] *Id.* at xix, xxiii.

[37] *Id.* at xxv-xxvi.

running.[38]

From 1920 to 1936, women's participation in athletics increased, partly due to the admission of women into the Olympic games.[39] Nevertheless, by 1970 the view that women were biologically incapable of engaging in strenuous physical activity still prevailed. For example, the philosopher Paul Weiss wrote in 1969 that women should be viewed as "truncated males" who could not participate in demanding physical activity.[40] Similarly, in 1967 the American Medical Association maintained that girls should refrain from contact sports because of their biological fragility.[41] By 1972, women's participation in athletics was still minimal. In 1971, only seven per cent of the participants in interscholastic high school sports were female.[42]

Since 1972, women's athletic participation has markedly increased, although it remains far below male participation. In 1979, more than one-fourth of all two-year colleges offered no women's sports.[43] By 1980, although women comprised thirty percent of all participants in intercollegiate athletic programs at four-year colleges, women received only one-sixth of college athletic scholarships.[44] In 1979, the average budget for men's athletics at four-year colleges and universities was five times the average women's athletics' budget.[45]

Prior to the passage of Title IX of the Education Amendments of 1972,[46] judicial notions of male and female roles prohibited courts from

---

[38] *Id.* at xxxi.

[39] *Id.* at xxviii.

[40] *See* Weiss, *Women Athletes*, in OUT OF THE BLEACHERS, *supra* note 33, at 65:

> It is part of our cultural heritage to make an effort to avoid having women maimed, disfigured, or hurt. That is one reason why they do not usually engage in and are not officially allowed to compete in such contact sports as boxing, wrestling, football, and rugby . . . . One way of dealing with these disparities between the athletic promise and achievements of men and women is to view women as truncated males . . . [T]he performances of males can be treated as a norm, with the women given handicaps in the shape of smaller and sometimes less dangerous or difficult tasks.

[41] "The female breasts and other organs can be injured seriously by a sudden blow. The danger of scars, broken teeth, or other results of injury probably are more a psychological hazard for girls than boys." Twin, *supra* note 33, at xxxiv (quoting Higdon & Higdon, *What Sports for Girls?*, 45 TODAY'S HEALTH 21 (Oct. 1967)).

[42] *See generally* NATIONAL ADVISORY COUNCIL ON WOMEN'S EDUCATIONAL PROGRAMS, TITLE IX: THE HALF FULL, HALF EMPTY GLASS 39-50 (Fall 1981) (reporting results of National Federation of State High School Associations' biennial survey).

[43] *Id.*

[44] *Id.*

[45] *Id.*

[46] 20 U.S.C. §§ 1681-1686 (1976). Title IX provides that "[no] person . . . shall,

seriously considering the claim that women should be permitted to participate in athletics on an equal basis with men. For example, a Connecticut court justified the exclusion of girls from a public school's all male cross-country and indoor track teams in 1971, on the eve of passage of Title IX, with the following statement:

> [W]ith boys vying with girls in cross country running and indoor track, the challenge to win, and the glory of achievement, at least for many boys, would lose incentive and become nullified. Athletic competition builds character in our boys, we do not need that kind of character in our girls, the women of tomorrow.[47]

Although Title IX is best known for its effect on women's athletic opportunities, suits brought under the statute alleging discrimination in athletics have fared poorly and are unlikely to do better in the future.[48]

---

on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance. . . ." *Id.* § 1681(a). For further discussion of Title IX and athletics, see Cox, *Intercollegiate Athletics and Title IX*, 46 GEO. WASH. L. REV. 34 (1977); Gaal, Dilorenzo & Evans, *HEW's Final "Policy Interpretation" of Title IX and Intercollegiate Athletics*, 6 J.C. & U.L. 345 (1979-1980); Jensen, *Title IX and Intercollegiate Athletics: HEW Gets Serious About Equality in Sports?*, 15 NEW ENG. L. REV. 573 (1980); Johnson, *The Evolution of Title IX: Prospects for Equality in Intercollegiate Athletics*, 11 GOLDEN GATE 759 (1981); Kadzielski, *Postsecondary Athletics in an Era of Equality: An Appraisal of the Effect of Title IX*, 5 J.C. & U.L. 123 (1978-1979); Koch, *Title IX and the NCAA*, 3 W. ST. U.L. REV. 250 (1976); Comment, *Women and Athletics: Toward a Physicality Perspective*, 5 HARV. WOMEN'S L.J. 121 (1982) [hereafter *Women and Athletics*]; Comment, *Termination of Federal Funding to School Athletic Programs Under Title IX of the Education Amendments of 1972*, 5 SAN FERN. V.L. REV. 417 (1977); Comment, *Sex Discrimination and Intercollegiate Athletics: Putting Some Muscle on Title IX*, 88 YALE L.J. 1254 (1979).

[47] Hollander v. Connecticut Interscholastic Athletic Conference, No. 124,927 (Conn. Super. Ct. New Haven County, Mar. 29, 1971), *appeal dismissed*, 164 Conn. 654, 295 A.2d 671 (1972) (mem.); *see also Women and Athletics*, *supra* note 46, at 128-31 (discussing historical exclusion of women from sports).

[48] Prior to 1984, most courts had found that athletic programs were not covered by Title IX because they did not receive direct federal funding. *See, e.g.*, University of Richmond v. Bell, 543 F. Supp. 321 (E.D. Va. 1982) (granting University's motion for summary judgment because no athletic programs receive direct federal financial assistance); Bennett v. West Texas State Univ., 525 F. Supp. 77 (N.D. Tex. 1981) (granting summary judgment for defendant because federal financial assistance to athletic program was indirect); Othen v. Ann Arbor School Bd., 507 F. Supp. 1376 (E.D. Mich. 1981) (denying plaintiff's petition for attorney's fees on the ground that plaintiff was unlikely to prevail because athletic program received no direct federal financial assistance). *But see* Haffer v. Temple Univ., 688 F.2d 14 (3d Cir. 1982) (finding general funding sufficient for athletic program to be covered under Title IX). After the Supreme Court's recent decision in Grove City College v. Bell, 465 U.S. 555 (1984) (general federal financial assistance is not sufficient to require compliance of a specific

Courts have continued to use stereotypes about the physical differences between all men and women[49] to limit women's athletic opportunities. For instance, two federal courts have relied on the following "expert" view of the differences between men's and women's physical abilities to justify sex-segregated physical activity for males and females:

> [M]en are taller than women; stronger than women by reason of their greater muscle mass; have larger hearts than women and a deeper breathing capacity, enabling them to utilize oxygen more efficiently than women; run faster, based upon the construction of the pelvic area, which, when women reach puberty, widens, causing the femur to bend outward, rendering the female incapable of running as efficiently as a male.[50]

Not only do courts exaggerate the differences between men and women, implying for instance that no woman could run as fast as any man, but they consider only those physical traits traditionally valued by men. Yet, as commentator Lyn Lemaire noted, "the importance of brute strength in many athletic activities is overrated. . . ."[51] As this Article will explore, the emphasis on male physical traits prevails in the selection of firefighters and police officers.

Today, most women over the age of twenty-five attended primary

---

program or activity with Title IX), it is unlikely that athletic programs will be found covered in suits brought under Title IX. Further expansion of women's athletic opportunities must therefore rely on voluntary efforts by educational authorities or an amendment to Title IX. Legislative proposals have been introduced to amend Titles VI and IX of the Civil Rights Act and § 504 of the Rehabilitation Act to permit enforcement of those statutes when an institution receives general funding that does not go directly to the allegedly discriminatory program. *See* proposed Civil Rights Restoration Act of 1985, S. 431, 99th Cong., 1st Sess. (1985) (introduced Feb. 7, 1985 by Senator Kennedy) and H.R. 700, 99th Cong., 1st Sess. (1985) (introduced Jan. 24, 1985 by Representative Hawkins). This amendment, if passed, would permit some governmental regulation of athletic activities that do not receive direct federal funding, although it would not permit the government to cut off that general funding as an enforcement remedy.

[49] Men and women, in fact, do not differ markedly in physical characteristics. "Data comparing the performances of a random sample of women and men will produce graphs of two largely overlapping bell-shaped curves, one for males and one for females." *Women and Athletics*, *supra* note 46, at 126 (citing Oglesby, *The Masculinity/Femininity Game*, in WOMEN AND SPORT: FROM MYTH TO REALITY 75, 78-79 (C. Oglesby ed. 1978)).

[50] Brenden v. Independent School Dist., 342 F. Supp. 1224, 1233 (D. Minn. 1972), *aff'd*, 477 F.2d 1292 (8th Cir. 1973) (citing Bucha v. Illinois High School Assoc., 351 F. Supp. 69 (N.D. Ill. 1972)).

[51] *Women and Athletics*, *supra* note 46, at 125 n.11; *see also* J. KAPLAN, WOMEN AND SPORTS 44 (1979) ("Strength is only part of the sporting equation . . . . To be effective in almost any sport, it's important to strike a balance between strength and flexibility.").

and secondary school before the implementation of Title IX in 1976.[52] They have had limited opportunity to develop their physical abilities. Physical examinations used as selection devices, especially on a rank-order basis, perpetuate exclusion of these women from a traditionally male domain — employment.[53] Since the opportunity to engage in physical activity has historically been a proxy for gender, these physical examinations must be viewed with suspicion.

B. Gender-Based Segregation in Employment

Historically, women were excluded from work outside the domestic sphere based on a conception of their natural biological temperament. The classic example of this conception is the 1872 case of *Bradwell v. Illinois*,[54] in which the Supreme Court upheld the state's denial of a woman's application to practice law. Speaking for a plurality of the Court, Justice Bradley justified this exclusion as fitting within women's biological role: "The natural and proper timidity and delicacy which belongs to the female sex evidently unfits it for many of the occupations of civil life."[55]

Thirty-five years later, the Supreme Court relied on considerations of women's physical abilities to justify an Oregon statute prohibiting the employment of women in any mechanical establishment, factory, or laundry for more than ten hours a day.[56] In the words of the Supreme Court:

> That woman's physical structure and the performance of maternal functions place her at a disadvantage in the struggle for subsistence is obvious. This is especially true when the burdens of motherhood are upon her. Even when they are not, by abundant testimony of the medical fraternity, continuance for a long time on her feet at work, repeating this from day to day, tends to injurious effects upon the body, and as healthy mothers are essential to vigorous offspring, the physical well-being of woman becomes an object of public interest and care in order to preserve the strength and vigor of the race.[57]

---

[52] Regulations to implement Title IX were not issued until 1975. 34 C.F.R. §§ 106-106.71 (1985) (HEW regulations on Nondiscrimination on the Basis of Sex in Education Programs and Activities Receiving or Benefiting from Federal Financial Assistance). These regulations provided that covered institutions had until one year from the effective date to comply with the regulations regarding athletics. *Id.* at § 106.9.

[53] *See infra* text accompanying notes 125-214.

[54] 83 U.S. 130 (1872).

[55] *Id.* at 141.

[56] Muller v. Oregon, 208 U.S. 412 (1908).

[57] *Id.* at 421.

In more recent years, the intertwining of biological arguments and discussions of women's fitness to perform certain occupations has become more subtle. An excellent example is *Dothard v. Rawlinson*,[58] in which the Supreme Court, while finding that minimum height and weight requirements violated Title VII, upheld an absolute exclusion of women from maximum security prison "contact" guard positions. The Court held that women would pose a substantial security problem, directly linked to their sex, and therefore that both female guards and prison security would benefit from the exclusion of women from contact positions.[59] In his dissent, Justice Marshall criticized the majority for relying on generalized statements about women's abilities. Moreover, he noted that the presence of women guards in Alabama prisons had led to no serious breaches of security.[60]

The history of women's employment in police and fire departments reflects these same stereotypes about women's proper role and biological abilities. Before 1972, many women were explicitly excluded from firefighting and law enforcement positions or were permitted only to work in certain "women's" positions.[61] For example, the City of Los Angeles maintained separate gender-based job classifications in entry-

---

[58] 433 U.S. 321 (1977).

[59] *Id.* at 334-37.

[60] *Id.* at 343 (Marshall, J., dissenting). According to Justice Marshall:

> [M]uch of the testimony of appellants' witnesses ignores individual differences among members of each sex and reads like "ancient canards about the proper role of women." The witnesses claimed that women guards are not strict disciplinarians; that they are physically less capable of protecting themselves and subduing unruly inmates; that inmates take advantage of them as they did their mothers, while male guards are strong father figures who easily maintain discipline, and so on. Yet the record shows that the presence of women guards has not led to a single incident amounting to a serious breach of security in any Alabama institution. And in any event, "Guards rarely enter the cell blocks and dormitories," where the danger of inmate attacks is the greatest.

Id. (citations omitted).

[61] *See, e.g.*, Blake v. City of Los Angeles, 595 F.2d 1367 (9th Cir. 1979), *cert. denied*, 446 U.S. 928 (1980) (women excluded from police work until 1973); Berkman v. City of New York, 536 F. Supp. 177, 182 (E.D.N.Y. 1982), *aff'd*, 705 F.2d 584 (2d Cir. 1983) (women excluded prior to 1972); United States v. City of Buffalo, 457 F. Supp. 612, 618 (W.D.N.Y. 1978), *aff'd as modified*, 633 F.2d 643 (2d Cir. 1980) (women excluded prior to 1973). This exclusion may have been unlawful prior to 1972 under the equal protection clause of the fourteenth amendment to the United States Constitution. *But see* Dothard v. Rawlinson, 433 U.S. 321 (1977) (upholding women's exclusion from contact guard positions for violent sex offenders in maximum security prison of state whose prison system had a history of rampant violence).

level police positions until 1973.[62] Women in the "policewoman" classification generally performed "tasks relating to women and children, desk duty, and administration."[63] Policewomen were barred from regular police patrol assignments and were ineligible for promotion above sergeant.[64] The gender-based stereotypes underlying this classification system are apparent from the types of tasks that women were permitted to perform. This practice of absolutely excluding women from police and firefighter positions generally ended around 1973[65] after Title VII became applicable to state and local governments on March 24, 1972.[66] Before 1972, women who were discriminated against in seeking employment in police or fire departments had no statutory remedy under federal law.

While abolishing absolute prohibitions against hiring women, many police and fire departments instituted minimum height and weight requirements that had a disparate impact upon women and certain racial minorities.[67] For example, the City of Los Angeles instituted a mini-

---

[62] *Blake*, 595 F.2d at 1371.

[63] *Id.*

[64] *Id.*

[65] *See supra* note 61.

[66] *See* § 706 of Title VII of the Civil Rights Act of 1964, as amended, 42 U.S.C. § 2000e-6 (1976 & Supp. 1979).

[67] *See, e.g.*, *Dothard*, 433 U.S. 321 (finding height and weight requirement for prison guards unlawful under Title VII); Costa v. Markey, 677 F.2d 158 (1st Cir. 1982) (invalidating height requirements for police officers used until November 1974); United States v. North Carolina, 512 F. Supp. 968 (E.D.N.C. 1981) (height requirement of five feet, six inches for state highway patrol officers operating to exclude more than three-fourths of female applicants); *City of Buffalo*, 457 F. Supp. at 625 (height requirements of five feet, seven inches for firefighters and five feet, nine inches for patrol officers, in use until 1974, having an adverse impact upon women and Spanish-surnamed Americans; composite height, weight, body frame, and physical fitness score, in use after 1974, having an adverse impact against women and Spanish-surnamed Americans); Guardians Ass'n v. Civil Serv. Comm'n, 431 F. Supp. 526 (S.D.N.Y. 1977), *aff'd in relevant part*, 630 F.2d 79 (2d Cir. 1980) (finding that five feet, seven inches minimum height requirement had a racially disproportionate impact against Hispanics and was not job related); League of United Latin Am. Citizens v. City of Santa Ana, 410 F. Supp. 873 (C.D. Cal. 1976) (invalidating five feet, seven inch height requirement for police officers which excluded approximately 50% of male Mexican-Americans); Officers for Justice v. Civil Serv. Comm'n, 395 F. Supp. 378 (N.D. Cal. 1975) (enjoining five feet, six inches height requirement for patrol officers for discriminating against Asians, Latins, and females); Hardy v. Stumpf, 37 Cal. App. 3d 958, 112 Cal. Rptr. 739 (1974) (striking down the height and weight requirements for Oakland police officers); State Div. of Human Rights v. New York City Dep't of Parks & Recreation, 38 A.D.2d 25, 326 N.Y.S.2d 178 (1971) (invalidating five feet, six inches in height and 125 pounds in weight requirement for lifeguard which excluded 90% of

mum height requirement of five feet, seven inches at the same time that it abolished sex-segregated job classifications.[68] Many employers discontinued using minimum height and weight requirements after 1977 when the Supreme Court held in *Dothard v. Rawlinson* that such requirements were unlawful under Title VII.[69]

In applying *Dothard*, courts have recognized the stereotypes about physical abilities underlying the use of minimum height and weight requirements. For example, in *Blake v. City of Los Angeles* the defendants attempted to justify their height requirement with the argument that taller officers could more easily control resisting suspects with minimum force and better observe field situations.[70] The Ninth Circuit rejected those unverified assumptions, observing that: "Shorter officers may have certain advantages in observing field situations (for example, the ability to look under objects, the ability to squeeze through narrow passageways) that taller officers lack."[71] Thus, the three judge panel recognized the importance of attributes other than those traditionally valued by men in police work.

In addition to instituting minimum height and weight requirements, many police and fire departments began to use physical performance examinations to select employees in the 1970's.[72] Some of these physical abilities tests have been successfully challenged under Title VII.[73] However, much litigation concerning these tests looms on the horizon.

---

women).

[68] *Blake*, 595 F.2d at 1371.

[69] 433 U.S. 321 (1977).

[70] 595 F.2d at 1379.

[71] *Id.* at 1379 n.7.

[72] *See, e.g.*, Harless v. Duck, 619 F.2d 611 (6th Cir. 1980) (police officers); *Berkman*, 536 F. Supp. 177; *Officers for Justice*, 395 F. Supp. at 381-82; Hardy v. Stumpf, 21 Cal. 3d 1, 576 P.2d 1342, 145 Cal. Rptr. 176 (police officers); Maine Human Rights Comm'n v. City of Auburn, 408 A.2d 1253 (Me. 1979) (city using height and weight requirements until 1975 for the position of police officer; using physical agility test since 1978); *cf.* Smith v. Olin Chem. Corp., 555 F.2d 1283 (5th Cir. 1977) (en banc) (court validating "good back" requirement for manual laborers without evidentiary hearing on validity despite test's adverse impact upon blacks and plaintiff's allegation that he could perform the job although he failed the physical examination). *But see* Corley v. City of Jacksonville, 506 F. Supp. 528, 533 n.9 (M.D. Fla. 1981) (noting that physical agility requirements were discontinued as the result of a validation study conducted by the Florida State Fire College that found them to be unrelated to job performance).

[73] *See infra* text accompanying notes 127-33.

C. Assumptions Underlying Gender-Based Discrimination Relating to Physical Activity.

Sex-based segregation of physical activity in athletics and employment has been justified with the stereotype that all physical differences between men and women are biologically destined. If the differences are biological, there is no point in encouraging women to be physically active on the same basis as men.[74] In fact, the segregation of physical activity on the basis of sex has exacerbated socially created, not biological, physical differences.[75] It has precluded women not only from developing physical skills on the athletic field but also from developing skills useful to some physically demanding occupations.[76] Despite the elimination of most overt job advertising on the basis of sex,[77] almost eighty percent of employed men work in professions that are more than eighty

---

[74] *See supra* text accompanying notes 29-73. *See generally* A. DWORKIN, WOMAN-HATING ch. 9 (1974) (arguing that even the sex distinction itself may be socially created); A. OAKLEY, SEX, GENDER & SOCIETY 30 (1972) (noting that sex differences are not uniform among all racial or ethnic groups; in some ethnic groups women are as tall as men, have equally broad shoulders and breasts as small as men's); Jaggar, *Human Biology in Feminist Theory: Sexual Equality Reconsidered*, in BEYOND DOMINATION: NEW PERSPECTIVES ON WOMEN AND PHILOSOPHY 41 (C. Gould ed. 1983) (arguing that equality between the sexes requires the transformation not only of so-called sex roles or of gender norms, but of those biological aspects of human nature that we have thought hitherto, in Firestone's words, as "the sex distinction itself"). For further discussion of women's athletic abilities, see generally OUT OF THE BLEACHERS, *supra* note 33; Rose, *The ERA and Women's Sport: An Hypothetical Trial Case*, in WOMEN AND SPORT: FROM MYTH TO REALITY, *supra* note 49, at 239-41.

[75] *See generally* OUT OF THE BLEACHERS, *supra* note 33.

[76] *See, e.g.*, Harless v. Duck, 619 F.2d 611 (6th Cir. 1980); Blake v. City of Los Angeles, 595 F.2d 1367 (9th Cir. 1979); Berkman v. City of New York, 536 F. Supp. 177 (E.D.N.Y. 1982).

[77] *See* U.S. DEPT. OF LABOR, BUREAU OF LABOR STATISTICS, EMPLOYMENT AND EARNINGS, Vol. 27, No. 1 (Jan. 1980). In 1979, men comprised 58.3% of the paid workforce, totaling 56,519,000 persons. More than 80% of the workers in the following job categories were male: managers (7,929,000); blue-collar workers (26,166,000); transport equipment operatives (3,319,000); nonfarm laborers (4,138,000); protective service workers (1,282,000); and farmworkers (2,216,000). These male employees comprised 4,505,000 persons or 79.7% of male workers. Women were 41.7% of the workforce, totaling 40,426,000 persons. Women comprised more than 80% of the workers in the following professions: librarians (152,000); nurses (1,386,000); pre-secondary school teachers (1,386,000); sales demonstrators or peddlers (334,000); clerical workers (14,143,000); dressmakers, sewers (876,000); private household service workers (1,062,000); lodging quarters cleaners (174,000); food counter workers, waiters, or food service workers (2,002,000); health service workers (1,643,000); and personal service workers (1,370,000). Those employees comprised more than 60% of female workers.

percent male.[78] Therefore, it is important to break down the employment barriers that perpetuate a sex-segregated workplace as part of a sex-segregated society.

Although sex-segregation continues in athletics, most employers have eliminated overt sex-based job selection. However, the exclusion of women from many traditionally male jobs continues, because of the institution of rank-order physical abilities examinations to select employees. Three assumptions underlie the development and use of physical abilities examinations; these assumptions often make the tests unlawful under Title VII. First, employers often assume that all traditionally male occupations and no traditionally female occupations require physical abilities selection devices.[79] They make these observations from the perceived "maleness" or "femaleness" of occupations rather than from a consideration of the occupation's actual physical requirements. For example, the traditionally female nursing profession[80] has never used height/weight requirements or physical performance examinations to select employees, despite the fact that nurses often must lift and carry patients. In contrast, firefighter applicants frequently are required to take physical performance examinations,[81] justified in part by the assertion that firefighters must carry victims from fires.[82] Title VII requires employers to justify the need for a physical abilities selection device that has a disparate impact upon a protected class by exploring whether it is necessary to screen employees based on physical ability.[83]

---

[78] *See* Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376 (1973).

[79] *See, e.g.*, *Harless*, 619 F.2d at 619 (employer testifying that test was developed through an "intuitive process"); *Blake*, 595 F.2d at 1378 (trial court approving a physical performance test for police officers with the justification that the validity of the examination was "obvious" without exploring the legal requirements for test validity); Vulcan Soc'y v. Civil Serv. Comm'n, 490 F.2d 387, 398 (2d Cir. 1973) (court implying that physical performance examination should have been used on a rank-order rather than qualifying basis). *See infra* text accompanying notes 125-214.

[80] *See supra* note 77. The author wishes to acknowledge the assistance of Mary Whisner in thinking of the example of the nursing profession. Other examples are librarians and child care workers.

[81] *See infra* text accompanying notes 125-214.

[82] *See, e.g.*, *Berkman*, 536 F. Supp. 177. In fact, it is not true that firefighters should carry victims from fires; the safest technique is to drag the victim so that he or she can avoid the rising smoke. The *Berkman* court recognized this fact in finding that the dummy carry test was invalid because "a carry of the type [over the shoulder] illustrated by the dummy carry portion [of the exam] would be dangerous in the extreme to carrier and carried alike." *Id.* at 207. *See infra* text accompanying notes 125-42.

[83] *See infra* text accompanying notes 125-42.

Second, employers often assume that the only physical demands of traditionally male activities involve the male-valued attributes of strength and speed.[84] Therefore, employers ignore the importance of physical attributes traditionally held by women such as flexibility, small-hand motor coordination, balance, endurance, and agility.[85] Moreover, stereotypes about the difficulty of attaining strength and speed[86] lead employers to ignore the significant component of skill, as opposed to strength, in many physically demanding tasks.[87] Title VII requires employers to identify the physical attributes a job requires rather than to assume that the only physical demands involve attributes most commonly held by men.[88]

Third, employers often assume that physical strength and speed should be valued on a rank-order basis.[89] They assume that "more is better" without objectively investigating whether an employee may need a minimum amount of physical strength or speed to perform a job as effectively as possible. Female applicants often have sufficient strength and speed to pass tests that measure the physical abilities needed to

---

[84] *See, e.g.*, *Harless*, 619 F.2d 611 (physical performance test of speed and strength for police officers); Association Against Discrimination v. City of Bridgeport, 594 F.2d 306, 311 (2d Cir. 1979) (court proposing that employer adopt a physical abilities test of speed and strength for firefighters); Association Against Discrimination v. Weeks, 454 F. Supp. 758, 759 (D. Conn. 1978) (fire department arguing that superior physical speed and strength are important traits for firefighters despite the fact that the hiring process only required applicants to complete a pass/fail agility test); *In re* Sontag v. Bronstein, 33 N.Y.2d 197, 351 N.Y.S.2d 389 (1973) (barbell test for position of audio-visual aid technician); *cf.* Attorney General v. Massachusetts Interscholastic Ass'n, 378 Mass. 342, 393 N.E.2d 284, 293 (1979) (defendants argued that differences between male and female athletic opportunities can be justified because "the discrimination under attack is in reality not one based on sex but rather on functional differences deriving in the main from biology").

[85] *See infra* text accompanying notes 168-85.

[86] Some employers purport to value and measure endurance. *See, e.g.*, *Berkman*, 536 F. Supp. 177. Because the test for endurance is generally short, the test is one of speed, not endurance. *See id.* at 196; *see also* W. MCARDLE, *supra* note 16, at 86. To avoid perpetuating the false perception that these employers' practices value endurance, this Article labels this characteristic "speed."

[87] *See, e.g.*, Officers for Justice v. Civil Serv. Comm'n, 395 F. Supp. 378, 382 n.1 (N.D. Cal. 1975) (defendants arguing physical agility test for police officer that had a complete exclusionary effect against women did not cause unlawful disparate impact against women because one would expect women to do less well than men on a physical agility test); Hardy v. Stumpf, 21 Cal. 3d 1, 6 n.1, 576 P.2d 1342, 1344 n.1, 145 Cal. Rptr. 176, 178 n.1 (1978) (ignoring fact that woman passed the physical performance test on her third attempt; however, only two attempts were permitted).

[88] *See infra* text accompanying notes 168-85.

[89] *See infra* text accompanying notes 186-214.

perform a job adequately. However, since females often do not have relatively more strength or speed than male applicants, the use of these examinations on a rank-order basis has a devastating impact on women's employment opportunities.[90] Title VII requires employers to provide a strong justification for the use of rank-order examinations that exclude a protected class, even if the examinations are valid on a cutoff basis.[91]

A discussion of rank-order physical abilities examinations to select employees in traditionally male occupations cannot operate from these three false premises. Instead, it must be grounded in an understanding of how these selection devices perpetuate gender-based stereotypes about physical abilities. For example, it is true that women, on average, have been deprived of the opportunity to attain many physical skills. However, women often can develop the physical skills needed for a particular job through a brief training program.[92] In fact, women, on average, improve more than men during a physical training program.[93] If one assumed that gender-based differences in physical abilities were not biologically destined but could be overcome, then the importance of pre-employment physical training programs for women would be more obvious.

II. STANDARDS FOR TEST VALIDATION UNDER TITLE VII

Extensive case law indicates that unvalidated rank-order written examinations that have an exclusionary impact on the employment opportunities of blacks and Hispanics are unlawful under Title VII of the Civil Rights Act of 1964.[94] The United States Department of Justice

---

[90] *Id.*

[91] *See infra* text accompanying notes 192-99.

[92] *See* testimony of McArdle and Magel in United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y Mar. 21-24, 1983) and in Berkman v. City of New York, 536 F. Supp. 177, 194 (E.D.N.Y. 1982).

[93] *See* Wilmore, *Alterations in Strength, Body Composition, and Anthropometric Measurement Consequent to a Ten-Week Weight Training Program*, 6 MED. SCI. SPORT 133 (1974) (showing women having a greater percent improvement than men in leg press, bench press, and grip strength after a ten-week training program); UNITED STATES ARMY RESEARCH INSTITUTE OF ENVIRONMENTAL MEDICINE, ANALYSIS OF ATTRITION, RETENTION AND CRITERION TASK PERFORMANCE OF RECRUITS DURING TRAINING 10 (Feb. 1982) (reporting that the percent discrepancies between men and women on various work capacity and strength measures were reduced following a seven-week training program) [hereafter ARMY REPORT].

[94] 42 U.S.C. §§ 2000e-2000e-17 (1972 & Supp. 1976).

has been on the forefront of developing much of this case law.[95] Unvalidated physical abilities examinations that limit women's opportunities for employment must be viewed analogously to unvalidated written examinations. Physical abilities examinations are as closely tied to the historical exclusion of women from physical activity[96] and "men's work"[97] as written examinations are tied to the historical exclusion of blacks from education[98] and "white work."[99] Nevertheless, neither courts nor the Department of Justice have applied this case law consistently to physical abilities tests that have an exclusionary impact upon women.[100]

Employment selection devices[101] are not lawful under Title VII

---

[95] *See, e.g.*, Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975) (United States arguing as amici curiae that pre-employment tests were invalid under Title VII); Ensley Branch of the NAACP v. Seibels, 616 F.2d 812 (5th Cir.), *cert. denied*, 449 U.S. 1061 (1980) (United States alleging that tests used by the personnel board to screen and rank applicants for positions as police officers and firefighters were unlawful); Walston v. County School Bd., 492 F.2d 919 (4th Cir. 1974) (United States alleging that the school board had instituted a qualification and selection plan for the hiring and retention of faculty members which had substantially reduced its black teaching staff); Vulcan Soc'y v. Civil Serv. Comm'n, 490 F.2d 387 (2d Cir. 1973) (United States recommending that court approve consent decree that would discontinue the use of written examination for firefighters having a disparate impact upon blacks and Hispanics); United States v. City of Buffalo, 457 F. Supp. 612 (W.D.N.Y. 1978), *aff'd as modified*, 633 F.2d 634 (2d Cir. 1980) (United States alleging that written examinations for police officer and firefighter that had a disproportionate impact against blacks were unlawful). *But see* United States v. Jefferson County, No. 75-P-0666-S (N.D. Ala. Feb. 27, 1984) (United States seeking to intervene on behalf of white plaintiffs alleging that written examinations administered by the personnel office should be used to appoint firefighters despite adverse impact of that selection device against blacks).

[96] *See supra* text accompanying notes 29-73.

[97] *See supra* note 77.

[98] *See, e.g.*, Brown v. Board of Educ., 347 U.S. 483 (1954); *see also* Brown v. Board of Educ., 349 U.S. 294 (1955) (requiring "all deliberate speed" in enforcing desegregation). *See generally* S. GOULD, THE MISMEASURE OF MAN (1981) (criticizing biologically deterministic view of sexual and racial cognitive differences). It is interesting to note that historically women, as well as blacks and other racial minorities, have been considered to have inferior cognitive capacities. *See id.*

[99] Written examinations have often been instituted soon after an employer began permitting blacks to apply for employment for the first time. *See* cases cited *supra* note 95.

[100] *See supra* text accompanying notes 11-19. *But see* United States' Post-Trial Brief with Respect to the Validity of Defendants' Selection Devices for Firefighters, United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y. filed Aug. 5, 1983) (arguing that rank-order use of physical performance for selecting firefighters is invalid and unlawful).

[101] *See* Uniform Guidelines, *supra* note 15; DIVISION OF INDUSTRIAL-ORGANIZA-

when they have an exclusionary impact[102] on the basis of race, sex, color, national origin, or religion unless employers can show them to be valid.[103] Under the Uniform Guidelines on Employee Selection Procedures,[104] the accepted regulations defining validation,[105] employers may validate selection procedures based on content, construct, or criterion-related validity studies. The appropriateness of the validity strategy depends upon the type of job for which the employer developed a selection device, the way in which the employer will use the test results, and available empirical data.

A content validity study consists of data showing that the "content of the selection procedure is representative of the important aspects of performance on the job for which the candidates are to be evaluated."[106] Content validation is appropriate when test items directly measure important job tasks that are a prerequisite to entry level job performance (for example, being able to type for a job that requires an experienced typist).[107] A proper content validation study must include a thorough job analysis identifying the important knowledge, skills, and abilities necessary to successful job performance.[108] Therefore, if an employer is

---

TIONAL PSYCHOLOGY, AMERICAN PSYCHOLOGICAL ASSOCIATION, PRINCIPLES FOR THE VALIDATION AND USE OF PERSONNEL SELECTION PROCEDURES (1980) [hereafter DIVISION 14 PRINCIPLES]; JOINT COMMITTEE OF APA, AERA & NCME, STANDARDS FOR EDUCATIONAL & PSYCHOLOGICAL TESTS (1974) [hereafter APA STANDARDS].

[102] *See supra* note 15 (defining exclusionary impact).

[103] *See supra* text accompanying notes 100-02; *infra* text accompanying notes 104-24.

[104] 29 C.F.R. §§ 1607-1607.18 (1985).

[105] As guidelines of the enforcement agencies, the Uniform Guidelines, *supra* note 15, are entitled to great deference. *Albemarle Paper Co.*, 422 U.S. 405; Griggs v. Duke Power Co., 401 U.S. 424 (1971). They should be followed absent some cogent reason to the contrary. Guardians Ass'n v. Civil Serv. Comm'n, 630 F.2d 79 (2d Cir. 1980), *cert. denied*, 452 U.S. 940 (1981); *Ensley Branch of NAACP*, 616 F.2d 812; Blake v. City of Los Angeles, 595 F.2d 1367 (9th Cir. 1979).

[106] Uniform Guidelines, *supra* note 15, § 1607.5B. The selection procedure in a content-valid test attempts to simulate the most important tasks in a job. By using a good job analysis and empirical data, a content validity study is far more than an attempt to attain "facial validity."

[107] *See supra* note 106. Whether a typing test is appropriate would depend on what skills are necessary for entry level typists. If they are trained or given an opportunity to learn to type on the job, a typing test that measures speed and accuracy may not be appropriate since those skills will improve with job experience. By contrast, if entry level typists are required to type with speed and accuracy immediately upon being hired, then a typing test may be appropriate.

[108] *Guardians Ass'n*, 633 F.2d at 242; APA STANDARDS, *supra note* 101, § E; DIVISION 14 PRINCIPLES, *supra* note 101, at 4, 13. This requirement is especially

seeking to justify a physical performance test, it is important that the job analysis focus on the job's physical requirements. In addition, a content validity study must investigate the skills that will readily be developed on the job because it is inappropriate to test for such skills.[109] For example, if an employer offers a physical training program for new employees, it is often difficult to validate the use of a physical examination for job applicants.

A construct validity study consists of data showing that "the procedure measures that degree to which candidates have identifiable characteristics which have been determined to be important in successful performance in the job for which candidates are to be evaluated."[110] Instead of directly measuring actual job tasks, construct-valid selection devices test for constructs that serve as a proxy for necessary job abilities.[111] Construct validation is appropriate when a job requires abstract qualities that cannot be measured directly, such as creativity.[112] Physical traits are not well suited for a construct validity strategy because they are observable rather than theoretical. Moreover, the concept of "task specificity"[113] recognized in the physical domain makes it very difficult to develop abstract indicators of physical performance. Valid physical tests must measure the exact muscle groups and level of intensity needed to perform a task. Rather than adhere to the requirement of task specificity, physical performance tests usually use outmoded "constructs" of "general fitness" such as pushups, pull-ups, and situps, although performance on these tasks has no demonstrated link to any job requirements.

A criterion-related validity study consists of "empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance."[114] Criterion vali-

---

important when an employer attempts to adopt a test developed in another jurisdiction. *See* Dickerson v. United States Steel, 472 F. Supp. 1304, 1338 (E.D. Pa. 1979) (failure to perform adequate job analysis is a "fatal flaw").

[109] *See* Uniform Guidelines, *supra* note 15, § 1607.5F (cautioning against selection on basis of knowledge, skills, or abilities that can be learned in brief orientation period).

[110] *Id.*

[111] *Id.*

[112] *See* APA STANDARDS, *supra* note 101, at 29, 46-47.

[113] "Task specificity" refers to the concept that the body's reaction to training is highly specific. It is important to train the specific muscles involved in a physical activity to create a training effect. For instance, there is a negligible "transfer" effect between developing aerobic power for swimming and for running. W. MCARDLE, *supra* note 16, at 268. *See also infra* text accompanying note 183.

[114] *See* APA STANDARDS, *supra* note 101. Criterion-related validity studies, like

dation is appropriate when a measure of job performance exists so that the correlation between test results and job performance can be examined.[115] Because of the difficulties of objectively measuring job performance, few test makers have attempted to use this validation technique to develop selection devices.

Irrespective of which validation strategy an employer chooses, valid employment selection devices must also be used in a manner that corresponds with the requirements of the job. Accordingly, tests may be valid for some uses (for example, cutoff basis) but not for others (for example, rank-order basis).[116] Hence, employers may not use physical performance tests on a rank-order basis without empirical evidence that "more is better."

In general, the use of an examination for ranking is more suspect than its use on a cutoff basis because ranking often causes extreme disparate impact.[117] Ranking must therefore satisfy the same legal standards as high cutoff scores that produce extreme disparate impact.[118] Ranking must select applicants based on an assessment of the minimum abilities needed for successful job performance.[119] Accordingly, incum-

---

construct validity studies, make use of empirical evidence. The difference between the two types of studies is that construct validity studies directly measure traits necessary to successful job performance whereas criterion validity studies may provide a more theoretical connection to successful job performance.

[115] *See supra* note 114.

[116] *See, e.g.*, Sarabria v. Toledo Police Patrolman's Ass'n, 601 F.2d 914 (6th Cir. 1979); Louisville Black Police Officers Org. v. City of Louisville, 20 Fair Empl. Prac. Cas. (BNA) 1195, 21 Empl. Prac. Dec. (CCH) ¶ 30,330 (W.D. Ky. 1979) (finding written police officer examination valid for establishing a cutoff but not for ranking); *cf. Ensley Branch of NAACP*, 616 F.2d 812 (finding completion of academy but not academy grades a valid criterion measure for validating a ranked, written examination).

[117] *See, e.g.*, *Guardians Ass'n*, 630 F.2d 79; *Sarabria*, 601 F.2d 914; *Vulcan Soc'y*, 490 F.2d 387; Vulcan Soc'y v. Fire Dep't, 505 F. Supp. 955 (S.D.N.Y. 1981); *Louisville Black Police Officers Org.*, 20 Fair Empl. Prac. Cas. (BNA) 1195; *see also* Uniform Guidelines, *supra* note 15, § 1607.5G ("Evidence which may be sufficient to support the use of a selection procedure on a pass/fail (screening) basis may be insufficient to support the use of the same procedure on a ranking basis under these guidelines.").

[118] *See supra* note 117.

[119] The Uniform Guidelines state:

> Where cut-off scores are used, they should normally be set so as to be reasonable and consistent with normal expectations of acceptable proficiency within the work force. Where applicants are ranked on the basis of properly validated selection procedures and those applicants scoring below a higher cut-off score than appropriate in light of such expectations have little or no chance of being selected for employment, the higher cut-off score may be appropriate, but the degree of adverse impact should be

bents must be able to pass a test for job applicants.[120] Administrative convenience (for example, the need to consider and hire a limited number of applicants) is not a sufficient justification for the establishment of norms; an independent, nonsubjective, and nonarbitrary basis for the cutoff score must exist.[121]

To validate a rank-order examination, a user must show that a higher score correlates with better job performance, even if the user has chosen a content-validation strategy.[122] Because content-validation studies do not usually gather that kind of evidence, this requirement often makes it impractical to use a content-validation approach to validate a test for use on a rank-order basis.[123] In sum, validation is required

---

considered.

Uniform Guidelines, *supra* note 15, § 1607.5H.

[120] *See Albemarle Paper Co.*, 422 U.S. 405; *Blake*, 595 F.2d 1367; Rogers v. International Paper Co., 510 F.2d 1340 (8th Cir. 1975) (60% pass rate by incumbents not sufficient); Officers for Justice v. Civil Serv. Comm'n, 395 F. Supp. 378 (N.D. Cal. 1975) (65% pass rate by incumbents not sufficient). *But see* Bridgeport Guardians v. Bridgeport Police Dep't, 431 F. Supp. 931 (D. Conn. 1977) (employer entitled to upgrade workforce); Hardy v. Stumpf, 21 Cal. 3d 1, 576 P.2d 1342, 145 Cal. Rptr. 176 (1978).

[121] *See, e.g., Guardians Ass'n*, 630 F.2d at 104; Shack v. Southworth, 521 F.2d 51 (6th Cir. 1975) (arbitrary passing score is equivalent to a subjective evaluation system); Walston v. County School Bd., 492 F.2d 919 (4th Cir. 1974), *aff'd & rev'd*, 566 F.2d 1201 (4th Cir. 1977); Corley v. City of Jacksonville, 506 F. Supp. 528, 536 n.12 (M.D. Fla. 1981) (cutoff scores for written firefighter examination upheld when set in reference to minimum knowledge, skills, and abilities necessary for the job); Kirkland v. New York State Dep't of Correctional Serv., 374 F. Supp. 1361, 1377 (S.D.N.Y. 1974), *rev'd on other grounds*, 520 F.2d 420 (2d Cir.), *reh'g denied*, 531 F.2d 5 (2d Cir. 1975), *cert. denied*, 429 U.S. 823 (1976) (selection of passing score invalid because it subordinated the goal of job relatedness to that of administrative convenience).

[122] The user must show "that a higher score on a content valid selection procedure is likely to result in better job performance. . . ." Uniform Guidelines, *supra* note 15, § 1607.14C(9).

[123] Nevertheless, the Questions and Answers interpreting the Uniform Guidelines, 44 Fed. Reg. 11,996 (1979), provide some guidance as to what kind of evidence might justify rank-order use of a content-validated physical examination. Question 62 asks, "Under what circumstances may a selection procedure be used for ranking?" The answer provides the following example:

> For example, for a particular warehouse worker job, the job analysis may show that lifting a 50-pound object is essential, but the job analysis does not show that lifting heavier objects is essential or would result in significantly better job performance. In this case a test of ability to lift 50 pounds could be justified on a content validity basis for a pass/fail determination. However, ranking of candidates on relative amount of weight that can be lifted would be inappropriate.

Id. at 12,005 (Answer 62).

when an examination has disparate impact on a protected class. Closer scrutiny is necessary when an employer uses the test on a rank-order basis and causes extreme disparate impact.[124] A relationship to job performance must usually be shown even when the test has been validated for cutoff use by a content-validation technique.

III. Application of Title VII Validity Standards to Rank-Order Physical Abilities Selection Devices

A. Assumption that Job Requires a Physical Abilities Selection Device

1. Prior Experience of Employer

Many employers have assumed that historically male jobs require physical abilities selection devices without investigating that assumption. They therefore have instituted physical performance examinations that disparately exclude women from employment. Although Title VII requires an employer to validate examinations having a disparate impact on a protected class,[125] most employers have failed to conduct even a cursory validation study of physical performance examinations.

Nevertheless, an employer's own employment experience may help invalidate the assumption that the job requires a preselection physical examination. An employer's own experience in successfully hiring employees without physical abilities tests — often before women were permitted to apply for employment — can show the lack of need for such a test.[126] Moreover, when viewed in this historical context, the use of a

---

The Questions and Answers also indicate empirical evidence is not always necessary to validate the rank-order use of a selection device. The relationship between scores and job performance may be inferred depending upon how "closely and completely the selection procedure approximates the important work behaviors." *Id.*

[124] However, public safety considerations may reduce the level of scrutiny. *See, e.g.*, Spurlock v. United Airlines, Inc., 475 F.2d 216, 219 (10th Cir. 1973) (requiring lesser justification for selection criteria because of the public safety involved in pilot's job performance). ***But see*** Vulcan Soc'y v. Fire Dep't, 505 F. Supp. 955, 965-66 (S.D.N.Y. 1981) (public safety rationale for firefighters insufficient to justify a high school diploma requirement).

[125] According to the Uniform Guidelines, "Any validity study should be based upon a review of information about the job for which the selection procedure is to be used. The review should include a job analysis except as provided in section 14B(3) below with respect to criterion-related validity." Uniform Guidelines, ***supra*** note 15, § 1607.14A. Under the provisions for content validity, the Guidelines state: "The work behavior(s) selected for measurement should be critical work behavior(s) and/or important work behavior(s) constituting most of the job." *Id.* at § 1607.14C(2).

[126] *See infra* notes 127-42.

physical abilities test may be a pretext for excluding women from jobs that have traditionally been considered "male."

Successful legal actions brought against the cities of Toledo, Ohio[127] and Los Angeles, California[128] for unlawful testing practices in the selection of police officers demonstrate the legal significance of tracing that historical relationship. The Sixth Circuit in *Harless v. Duck*[129] held that a physical examination was not a necessary selection device for the Toledo Police Department.[130] The court relied in part on the fact that the Department had deleted the physical ability test from a more recent selection procedure with no detrimental effect.[131] Similarly,

---

[127] *See* Harless v. Duck, 619 F.2d 611 (6th Cir. 1980).

[128] *See* Blake v. City of Los Angeles, 595 F.2d 1367 (9th Cir. 1979).

[129] 619 F.2d at 614.

[130] *Id.* at 617.

[131] *Id.* From 1965 to 1975, only 12 women were employed by the Toledo Police Department (the TPD). In 1975, the TPD employed 717 male police officers, one female police officer, and four policewomen. Until November 1974, it maintained separate entry level classifications for male and female officers. Through 1972, the TPD's job announcement for police officer positions sought "able ambitious young men," although it accepted applications from 18 women. The Department also used a physical abilities examination until 1975 that excluded disproportionate numbers of women from employment. Only one of the 96 persons hired under these procedures in 1972 was female. *Id.* at 615. The court does not provide specific figures as to the exclusionary impact of the physical abilities examination. It does, however, provide overall figures for the exclusionary nature of the entire hiring process, which consisted of a physical abilities examination, a written examination, and a structured oral interview. The physical ability test had four parts, of which the applicant was required to pass three parts: 15 push-ups, 25 situps, a six foot standing broad jump, and a 25-second obstacle course. Nearly 80% (14 of 18) of the female applicants and fewer than 25% (64 of 263) of the male applicants failed to be placed on the eligibility list. *Id.* at 614-15.

In 1975, after the appointment of a new police chief, the TPD abolished the separate entry level classifications and began to use a test which had no statistically significant disparate impact on women. *Id.* at 614 n.2, 615. Twenty-two per cent of the applicants were female and 19.4% of the persons hired were female. *Id.* at 615. The percentage of women hired rose from 1% in 1972 to 19.4% in 1975. The Department found that it could successfully hire police officers without using an exclusionary physical abilities test.

Female applicants challenged the Toledo Police Department's former hiring practices as violating Title VII, the fourteenth amendment, and the Civil Rights Act of 1871 in *Harless*. The TPD sought to justify its physical abilities test with the argument that it was necessary as a replacement for the recently abolished height and weight requirements. *Id.* at 616-17. The Sixth Circuit rejected that argument and held, contrary to the district court, that the 1972 test was not job related according to Title VII standards. *Id.*

in *Blake v. City of Los Angeles*,[132] the Ninth Circuit held invalid the use of pre-employment physical tests partly because the Los Angeles Police Department had successfully hired thousands of male police officers without physical tests.[133]

---

[132] 595 F.2d 1371. Prior to July 1, 1973, the Los Angeles Police Department (LAPD) maintained separate, gender-based job classifications in entry level police positions and selected police officers without use of a physical abilities test. *Id.*

Between 1970 and 1973, no women were appointed to sworn positions in the LAPD, although the Department hired more than 2,000 men. *Id.* Beginning on July 1, 1973, the LAPD abandoned sex-segregated job classifications. At the same time it imposed a five feet, seven inch height requirement which was then lowered to five feet, six inches. The 1973 Los Angeles physical abilities test consisted of five events: running 50 yards and scaling a smooth wall six feet high; running 50 yards and hanging from a chinning bar using an overhand grip for one minute; running 50 feet and dragging a dead weight of 140 pounds for 50 feet; running 50 yards and holding a stylus steady for 17 seconds; running as many laps around a one-eighth mile track as possible in 12 minutes. *Id.* at 1381 n.14. It also instituted a physical abilities test. *Id.* at 1371. Although women were theoretically eligible for employment in all officer positions after 1973, the selection procedures maintained the prior record of exclusion. In 1976, women constituted only two percent of all sworn positions, approximately the same percentage that the Department had employed in 1970. *Id.*

Unsuccessful female job applicants brought suit against the City of Los Angeles Police Department in *Blake* alleging that the LAPD's selection procedures violated Title VII and related constitutional and statutory provisions. The LAPD sought to demonstrate the job relatedness of the physical abilities test through the results of two validation studies. *Id.* at 1381. The court of appeals reversed the district court's grant of summary judgment for the defendants, finding that defendants did not meet their burden of justifying the use of the physical abilities test as a business necessity. *Id.* at 1382-83.

[133] *Id.* at 1382. The court found this fact as well as evidence that other employers successfully hired employees without the use of a physical ability examination demonstrated that the physical ability requirement was not "essential to safe and efficient job performance," *id.* at 1382 n.15, under the following provision of the Uniform Guidelines:

> Those employees or applicants who have been denied equal treatment because of prior discriminatory practices or policies, must at least be afforded the same opportunities as had existed for other employees or applicants during the period of discrimination. Thus, the persons who were in the class of persons discriminated against during the period the user followed the discriminatory practices should be allowed the opportunity to qualify under less stringent selection procedures previously followed, unless the user demonstrates that the increased standards are required by business necessity.

Uniform Guidelines, *supra* note 15, § 1607.11. *Cf.* James v. Stockham Valves & Fitting Co., 559 F.2d 310, 337 n.43 (5th Cir. 1977), *cert. denied*, 434 U.S. 1034 (1978) (approving of but not relying on that interpretation of the Guidelines). *Contra* Hardy v. Stumpf, 21 Cal. 3d 1, 10, 576 P.2d 1342, 1346-47, 145 Cal. Rptr. 176, 179 (1978)

In retrospect, both *Harless* and *Blake* may seem to have been easy cases because of evidence that the employer had successfully hired police officers without use of a physical abilities examination, and that the examinations had only been instituted after women were permitted to seek employment. Nevertheless, the plaintiffs were unsuccessful in the district courts in both cases.[134] That history suggests that some lower courts may still be reluctant to recognize the inappropriateness of physical abilities requirements for police officers even with evidence of the historical exclusion of the protected class and the availability of alternative nondiscriminatory selection devices.

Although physical abilities tests are diminishing in the selection of police officers due to decisions like *Harless* and *Blake*, their use is growing in the selection of firefighters.[135] However, the history of firefighting examinations is remarkably similar to that of police tests. For example, prior to 1977, the City of Buffalo had an absolute prohibition against hiring women as firefighters[136] and did not use a physical abilities test to select firefighters.[137] Under court order, it began to accept applications from women in 1978.[138] At the same time Buffalo adopted a physical abilities test[139] and a new written examination.[140]

(finding that applicants for police officer could be required to pass a standard not imposed on current police officers; rejecting the relevance of 29 C.F.R. § 1607.11 to that issue).

[134] In *Blake*, the district court granted summary judgment to the city, finding that the plaintiffs did not meet their burden of proving a prima facie case. *Blake*, 595 F.2d at 1372. In *Harless*, the district court held that defendants did not violate Title VII since it found that all discrimination had ceased as of March 24, 1972. *Harless*, 619 F.2d at 615. Although it found that defendants had intentionally discriminated prior to March 1972, which violated the fourteenth amendment and 42 U.S.C. § 1983, it found that plaintiffs had not established any injury for which they were entitled to relief during that period. *Id.* at 615, 617.

[135] *See supra* note 17.

[136] United States v. City of Buffalo, 457 F. Supp. 612, 630 (W.D.N.Y. 1978), *aff'd as modified*, 633 F.2d 643 (1980).

[137] *Id.* at 625.

[138] *Id.* at 640.

[139] *See* United States' Pretrial Brief with Respect to the Validity of Defendants' Selection Devices for Firefighters, United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y. filed Mar. 17, 1983) at 12-19 [hereafter Buffalo Pretrial Brief]. Buffalo borrowed the test from the City of Chicago. The Buffalo and Chicago test consisted of a stair climb, flexed-arm hang, hose couple exercise, man lift and carry, and obstacle run. Each of these elements was timed. Chicago hired applicants based on their rank-order score on the physical abilities test if they passed a written examination. Buffalo hired applicants based on the average of their score on this examination and a written examination. *Id. See also* J. MISNER, R. BOILEAU & W. CONSIDINE, DEVELOPMENT

Not one woman has been hired as the result of these new procedures.[141] The only women hired in Buffalo were hired as the result of a court order.[142]

Therefore, the introduction of physical abilities tests after ending women's absolute exclusion must be viewed with suspicion. These tests often serve to perpetuate a policy of absolute exclusion. Moreover, as this Article shows, they are often introduced without any systematic study of the physical characteristics necessary for firefighting.

2. Failure to Investigate Whether Job is Physically Demanding

Most job analysis devices focus solely on cognitive job components; therefore, their usefulness in the physical domain is limited.[143] Few attempts have been made to develop job analysis studies that are sensitive to physical job requirements. Instead, "facial" observations[144] of job requirements are usually the major focus of physical job analyses.[145] Because of the concept of task specificity,[146] the requirement of assessing the physical demands of a job is especially important. Not only must the job's overall physical demands be assessed, but the muscle groups used and the required level of physical intensity must be studied as well.[147]

In *Harless*,[148] the Sixth Circuit recognized that a valid job analysis for a physical abilities test must focus on the physical demands of the job. The court found that the Toledo Police Department had never conducted a job analysis to determine the amount of physical strength or

---

OF A PHYSICAL PERFORMANCE TEST FOR CHICAGO FIREFIGHTER APPLICANTS — FINAL REPORT (1977) [hereafter FINAL REPORT].

[140] Buffalo Pretrial Brief, *supra* note 139, at 9-12. In 1977 Buffalo administered a written examination prepared by the New York State Civil Service Commission. In 1980, Buffalo administered a written examination, the IPMA B1(M), prepared by the International Personnel Management Association.

[141] Buffalo Pretrial Brief, *supra* note 139, at 16-19.

[142] On April 30, 1980, the court entered an order requiring the city to hire five female firefighters. *Id.* at 5.

[143] United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y. hearing held Mar. 21-25, 1983).

[144] Subjective observations of lay people are one example.

[145] *See* United States' Memorandum Concerning Jurisdiction and Alternative Selection Procedures, Hammon v. Barry, No. 84-0903 (D.D.C. filed Apr. 25, 1985).

[146] Because the job analysis must study the "important" or "critical" job components, it must study the physical components if it is to conclude that a physical abilities test is necessary. *See supra* notes 106-11.

[147] *See infra* text accompanying notes 157-67.

[148] 619 F.2d at 611.

extent of physical exertion required for the job; instead, as the court found, the sole justification for the examination was the wholly inadequate "intuition" of department officials.[149]

In the firefighting context, the district court in *Berkman v. City of New York*[150] recognized the invalidity of a job analysis for a physical performance test[151] that did not closely examine physical job demands. The city relied on a formal job analysis, facial observations, and advice from the City of Chicago to justify its examination.[152] The *Berkman* court found that New York's formal job analyses were too general and arbitrary,[153] and that the city compounded these errors by relying on facial observations by officials of the fire department and personnel office who knew nothing of the physical requirements of the job.[154]

The development of the time standard for the one mile run reflected

---

[149] *See id.* at 616.

[150] 536 F. Supp. at 177 (E.D.N.Y. 1982).

[151] The New York City examination included an agility test consisting of an obstacle course with two walls, one five feet and one eight feet, and a window-ledge-window obstacle; victim rescue, carrying a 120-pound dummy up and down one flight of stairs; ledgewalk; mile run; flexed-arm hang; hand grip; and broad jump. *Id.* at 197.

[152] *Id.* at 212-16.

[153] A physical abilities analysis (PAA), job inventory, critical incidents analysis, position analysis questionnaire, and physical demands analysis (PDA) formed a central part of the job analysis eventually relied upon to develop the firefighters' examination. The PAA attempted to use factors of physical performance developed by Edwin Fleishman, an industrial psychologist. Dr. Fleishman and his colleagues had developed a system of rating scales, known as the abilities analysis, which could be used to indicate on a scale of one to seven how much of a given ability is needed to perform a particular task. Until adopted by the City of New York, the test had never been used to analyze an entire job rather than a specific task. The court first assessed the validity of all the job analysis studies except the PAA. Although the court found that the job inventory was "by far the most far reaching inquiry made by [the consultants]," *id.* at 185 n.2, it found that the consultants did not analyze those results "in terms of physical abilities involved in the activity or in terms of the tests administered." *Id.*

[154] According to the court, the defendants "engaged in ex post facto rationalization of tests selected for quite different purposes, drawing on aspects of [the consultant's] job analysis which were never intended to be put to such use." *Id.* The court also found that the critical incidents analysis failed to include any systematic effort to determine whether the respondents understood the abilities being discussed and failed to assess the relative importance of the specific job behaviors reported in relation to the overall job and in relation to each other. Finally, the court found that the position analysis questionnaire was at a high level of generality and contradicted the results of the PAA, which itself had serious limitations. The court found that the PAA and its ranking of abilities were not so well grounded in the observable behaviors of firefighters that one could predict with any confidence the capacity of an individual to perform well as a firefighter. *Id.* at 190.

this problem. To pass this section of the examination and receive some points, an applicant had to complete the run in less than seven minutes, thirty seconds; to receive full credit for this item an applicant had to finish in less than four minutes, fifty seconds.[155] Based on the results of a trial run by job incumbents, the consultants had originally suggested that applicants who completed the run in less than twelve minutes pass this section of the examination. However, they agreed to reexamine the scoring plan to make it more stringent because

> both the Chief of the Fire Department and the Chief of the Department of Personnel complained of the scoring plan for the mile run (which permitted a candidate to complete the mile in 12 minutes and still not fail), not on the basis of its methodology, but simply because they or some close relative could run the mile in less time.[156]

The consultants reexamined the scoring plan, but they did not rely on the existing information about job incumbents or retest the job incumbents. Instead they referred to mile run scores of an all-male sample population involved in a program of daily physical training in preparation for military service in World War II.[157]

The City of Chicago has also used a faulty job analysis in developing its physical performance test. Buffalo and Chicago use the Chicago job analysis and the United States recommended that the District of Columbia fire department rely on Chicago's expertise in developing its own examination. Therefore, the limitations of the Chicago job analysis are important. The *Berkman* court recognized some of these problems when it criticized New York City for relying on advice from the City of Chicago to modify several test items.[158] The United States has challenged the Chicago job analysis in its litigation against the City of Buffalo, because Buffalo relied on the Chicago job analysis in developing its own physical performance examination.[159]

---

[155] *Id.* at 195-96.

[156] *Id.* at 196, 204.

[157] *Id.* at 196. The court found that the consultants "accepted the suggestion that this more stringent scoring norm be used with little if any attention to the logic of using such norms in its overall scoring system, either in terms of job requirements or norms in the applicant pool." *Id.* at 196. Thus, the city relied on unfounded stereotypes about the job's physical requirements to impose a standard derived from an irrelevant source even though it had available internal data showing the actual average time in which job incumbents were able to complete the mile run. Similarly, the city modified three other tests proposed by the consultants. The court found that "there appears no basis for saying that any of the three tests reliably measured the abstract abilities measured by the Fleishman tests for which they were substituted." *Id.* at 197.

[158] *Id.* at 197; *see also* Buffalo Pretrial Brief, *supra* note 139, at 20-30.

[159] The Chicago test is virtually identical to the test used by the City of Buffalo. *See*

Chicago used the functional job analysis system (FJA) to conduct its analysis.[160] Experts, including FJA's developer, Sidney Fine, commonly recognize that system as inappropriate for analyzing the physical demands of a job.[161] None of the categories of the FJA even refer to physical demands.[162] Despite a paucity of physical demand descriptions, the Chicago personnel board used the job analysis to divide the job of firefighter into physical components.[163]

Although the Chicago job analysis did not analyze physical job requirements, the test developers summarily concluded in their final report that the "great majority" of the tasks were physical in nature and broke them down into the following four categories: (1) manipulation,

---

supra note 139.

[160] *See* FINAL REPORT, *supra* note 139, Appendix A (Task Analysis Report).

[161] *See* Trial Transcript of Testimony of D. Gebhardt, United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y. Mar. 24, 1983). Gebhardt testified that she had worked with Sidney Fine for the last several years and had used his functional job analysis instrument. She then stated, "I believe that Sid would also say that one of the weaknesses of a functional job analysis is the identification of the physical demands in the job . . . . None of [the job analysis components] identifies any of the physical demands." *Id.* at 785-86.

[162] *See* FINAL REPORT, *supra* note 139, Appendix A.

[163] The job analysts determined that firefighters performed 81 tasks. They analyzed and rated those tasks using the following scales: worker functions, GED scale, worker instructions, and SVP. The complexity of each worker function was rated, with the data scale ranging from zero to six, the people scale from zero to eight, and the things scale from zero to seven. The GED scale showed the level of reasoning, math, and language development necessary to perform the task. The scale ranged from one to six. The worker instruction scale showed the level of discretion that the employee exercises in performing the task. It ranged from one to eight. The SVP scale estimated the amount of time necessary before the person is performing the task efficiently. It ranged from one to nine. None of these scales directly referred to physical demands. The closest any scale came to referring to physical demands was the "thing" scale. Of the 81 task statements that were rated, only 32 (40%) were rated as more than 50% "thing" oriented. *Id.*

The tasks were then broken down into two groups: high (worker function, worker instruction, and SVP) and low (worker function, worker instruction, and SVP). Worker functions determined whether the tasks were primarily oriented to data (mental), people, or things, and assigned the appropriate percentage breakdown. The tasks were divided by whether they were performed regularly or only occasionally. Only 10 of 81 (12%) tasks that were performed on a regular basis were rated as "high." The tasks were then reviewed to determine whether they were data (mental), people, or things oriented. Of the 81 tasks, 39.8% were defined as "data" oriented and 43.9% as "things" oriented, suggesting that the cognitive aspect of the job was as least as high as the physical dimension (even assuming that the "things" dimension was physical). Finally, the tasks were rated as to complexity for data, people, and things. *Id.*

(2) pushing-pulling, (3) body movements (climbing, crawling, etc.), and (4) physical strength — (a) under twenty-five pounds, (b) twenty-five to fifty pounds, and (c) over fifty pounds.[164] The final report did not discuss how developers made these assessments. Moreover, these physical estimates did not purport to analyze which muscle groups were used. Finally, no investigation of aerobic physical exercise was undertaken; all the physical activities were anaerobic in nature.[165] The highest level of physical strength considered was fifty pounds.[166] On the basis of this "job analysis" the test preparers concluded that the job was sixty percent physical in nature and that the physical exam should therefore be rated sixty percent.[167] The Chicago consultants thus developed a test that was primarily physical in nature based on no close analysis of the job's physical demands.

Despite extensive Title VII case law requiring analysis of the important abilities needed for successful job performance, cities have developed physical abilities tests for police officers and firefighters without examining the physical demands of the jobs. Such tests do not meet the validity requirements and have an adverse impact on the basis of sex.

B. Assumption that Strength and Speed Are the Only Physical Demands

Another common assumption is that, even if a job is considered physical, the only physical requirements involve speed and strength.[168] The few job analysis studies conducted for traditionally male employment are so cursory that they do not isolate the needed physical attributes.[169] Physical attributes women often possess that are important to successful job performance, such as coordination, agility, flexibility, or endurance, are disregarded.[170]

---

[164] *Id.*

[165] *Id.* An activity that lasts less than two minutes at maximal physical output is more than 50% anaerobic. W. MCARDLE, *supra* note 16, at 86. Because each of the test items took less than two minutes to complete, and applicants were permitted a rest period between each activity, the test was therefore composed of primarily anaerobic activities. At the hearing on the validity of the physical performance examination in United States v. Buffalo, No. 74 Civ. 195 (Mar. 24, 1983), Dr. Misner, a developer of the Chicago test, agreed with this characterization of the physical performance test as anaerobic. Trial transcript at 306 (Mar. 22, 1985).

[166] FINAL REPORT, *supra* note 139, Appendix A.

[167] *Id.*

[168] *See infra* text accompanying notes 171-82.

[169] *See id.*

[170] *See id.*

For instance, the test development consultants for New York City had identified the importance of flexibility and balance in firefighting and therefore recommended a twist-and-touch test and a balance test. City officials rejected those test items complaining that they wanted a test "for firefighters, not for ballet dancers."[171] They also relied on advice from Chicago officials who thought that these two tests were "inappropriate" for firefighters.[172] The city then substituted new test items for those two tests. In *Berkman v. City of New York* the court found the new items did not measure as accurately as the tests for which they had been substituted.[173] In addition, the new tests measured speed rather than flexibility and balance.[174]

Other evidence of the invalid emphasis on speed was the scoring of the mile run. The consultants claimed to be using that test item to measure aerobic capacity or endurance;[175] however, they in fact tested for anaerobic capacity (strength and speed) by requiring applicants to run at top speed to score well. In the court's words: "In the absence of pacing, the test measures not stamina or aerobic capacity . . . but anaerobic capacity — that is, the ability of the body to generate energy from itself, something also measured by strength tests."[176]

---

[171] Berkman v. City of New York, 536 F. Supp. 177, 196 (E.D.N.Y. 1982), *aff'd*, 705 F.2d 584 (2d Cir. 1983).

[172] *Id.* at 197.

[173] *Id.* at 208-16.

[174] *Id.* at 197. The city's gender bias is apparent from the ways it modified the tests of balance and flexibility. The City modified the flexibility test to score applicants based on how fast they completed the test item rather than by how far they could extend themselves. *Id.* This made the test one of speed rather than flexibility although, as the court found, "firefighting is not a sprint event." *Id.* at 204. Similarly, the city substituted a ledge walk for the balance test. *Id.* at 197. That test item also measured the applicants' speed rather than balance because they were scored based on their completion time and were permitted to perform the test with their eyes open. *Id.* The court rejected the validity of that test item because of the absence of "any factual inquiry as to the existence and importance in firefighting of having to side step at high speed forward and back over a narrow ledge." *Id.* Additionally, the court criticized the test's reliability as a measure of balance because applicants were allowed to open their eyes during the test.

Recently, the district court in New York examined the physical exam developed by the City of New York after *Berkman*. The court held that, although the exam tested important physical capacities necessary to the job, the city still placed too much importance on speed. The exam's scoring system gave speed and aerobic fitness undue emphasis, said the court. *See* Berkman v. City of New York, No. 79 Civ. 1813, 38 Empl. Prac. Dec. (CCH) ¶ 35,687 (E.D.N.Y. Oct. 7, 1985).

[175] *Id.* at 204.

[176] *Id.*

This emphasis on speed rather than endurance is typical of most firefighting tests, which consist of timed, sprint events.[177] A more appropriate test for endurance, as noted by the *Berkman* court, is a test requiring the candidate to run or walk for twelve minutes, measuring the distance accomplished.[178] Not only would such a test more accurately measure the job requirements, but it would have less adverse impact upon women than a test of speed.[179]

The emphasis on speed is linked to the emphasis on strength. One reason the *Berkman* court rejected the "sprint" test items was that they overemphasized strength as well as speed. The court found this emphasis inappropriate "because of the general consensus among the witnesses, both experts and job incumbents, that stamina rather than brute strength is the prime requirement for the job of firefighting."[180]

In addition, the *Berkman* court addressed the stereotype that all tests for strength or endurance are equally acceptable. The New York City consultants had relied on general constructs of physical abilities (for example, dynamic strength) to develop their test items.[181] The *Berkman* court rejected the use of general constructs because they did not have a "grounding in observable behavior of the way firefighters operate that one can say with confidence that a person who possesses a high degree of these abilities as opposed to others will perform well on the job."[182]

The court's unwillingness to permit defendants to use general constructs without specific evidence of job requirements is tied to the concept of "task specificity."[183] There is a negligible "transfer" effect between improvement in one physical ability and a seemingly closely

---

[177] Both the Chicago and Buffalo tests consist of timed, sprint events. *See supra* note 139.

[178] 536 F. Supp. at 204.

[179] *See* W. MCARDLE, *supra* note 16, at 135, 145.

[180] *Berkman*, 536 F. Supp. at 204.

[181] *Id.* at 186-88.

[182] *Id.* at 206.

[183] According to Dr. William McArdle, an expert in exercise physiology who testified at the *Berkman* trial, task specificity

> refers to adaptation in the *metabolic* and *physiologic* systems depending on the type of overload imposed. It is known that a specific exercise stress such as strength-power training induces specific strength-power adaptations, and that specific aerobic or cardiovascular exercise elicits specific endurance-training adaptations. The specificity principle, however, goes beyond this because development of aerobic fitness for swimming, bicycling, or running is most effectively achieved when the exerciser trains the specific muscles involved in that desired performance.

W. MCARDLE, *supra* note 16, at 268.

linked other physical ability (for example, endurance in swimming and endurance in running).[184] Thus, it was inappropriate for New York City to substitute one measure of strength for another measure of strength without evidence that the specific test item related to actual work tasks.[185]

In conclusion, jurisdictions have chosen test items that emphasize traditionally male-valued speed and strength, ignoring actual job requirements. When a job analysis in New York revealed that traditionally female-valued traits are useful, those results were ignored. Valid physical performance tests must examine all of a job's physical requirements, not solely the requirements of speed and strength.

C. *Assumption that Rank-Order Use of Physical Abilities Tests is Appropriate*

The stereotype that "more strength is better" has led employers to use physical performance tests on a rank-order basis and thereby exclude women from employment opportunities.[186] That stereotype is not necessarily valid. Moreover, extensive case law on rank-order written examinations requires employers to provide strong justification for rank-order tests having a nearly absolute exclusionary impact on a protected class.

Prior to women's entry into firefighting, courts rejected the validity of written examinations on a rank-order basis for selecting firefighters, recognizing that no evidence existed that "more is better." For instance, in *Association Against Discrimination on Employment, Inc. v. City of Bridgeport*,[187] the court found that competitive written civil service ex-

---

[184] *Id.* at 268-69 (citing Magel, *Specificity of Swim Training on Maximum Oxygen Uptake*, 38 J. APPL. PHYSIOL. 151 (1975)).

[185] For instance, New York City's consultants had originally recommended the use of push-ups as one test item to measure upper body strength. *Berkman*, 536 F. Supp. at 197. Based on the advice of the Chicago experts, New York City substituted the flexed-arm hang, *id.* at 197, although their own consultants had previously concluded that that test item was a relatively unreliable test of dynamic strength. *Id.* The test item had a low (.53) test/retest reliability attributable to difficulties in obtaining objective measurements because of subjective differences between monitors in deciding when the arms are completely extended. *Id.* at 202. Moreover, the city chose a scoring system that would have resulted in a large number of the incumbent New York City firefighters scoring at or below zero. *Id.* at 198. For these reasons, the court rejected the validity of that test item. *Id.* at 198-202.

[186] *See infra* text accompanying notes 187-214.

[187] 454 F. Supp. 751 (D. Conn. 1978), *aff'd in relevant part*, 647 F.2d 256 (2d Cir. 1981), *cert. denied.*, 455 U.S. 988 (1982).

ams are "inadequate to the task of predicting complex patterns of behavior (for example, those behaviors which would distinguish an average firefighter from an exceptional firefighter)."[188] Therefore, the court strongly recommended that the City of Bridgeport lower its passing score by half to eliminate the racial disparate impact.[189] The *Berkman* court also rejected the use of a cutoff score and ranking for a physical performance examination. It recognized that the test was used with a cutoff score and rank-ordering because it "was customary in competitive civil service testing."[190] Relying on the Second Circuit's holding in *Guardians Association v. Civil Service Commission*,[191] the court noted that rank-ordering can only be used if it can be shown that a higher score is likely to result in better job performance.[192] The city in *Berkman* justified rank-order use with the assertion that "every increment in the abilities tested for in its physical exam necessarily represents a better performance as a firefighter."[193] Its evidence was that the relationship was "obvious," without recourse to any empirical data.[194]

The court then rejected the defendant's assertions about the "obvious" relationship between higher scores and better performance:

> [W]hat must be identified are not those who are strongest or fastest but, instead, those who, with the benefit of training in pacing or because of their native capacities of endurance, can perform the punishing tasks of fire fighting as they are actually required to be performed. . . .
>
> Not only does maximal performance compromise the fire fighter's ability to pace performance over the long periods of time during which physical demands are being made on the body, in addition, the very unpredictability of fire and the instability of burning structures call upon qualities of foresight, endurance and pacing not examined by tests of maximum physical strength.[195]

Although the *Berkman* court soundly rejected the notion that "more is necessarily better," that assumption still underlies the use of many firefighting selection devices. For instance, the City of Chicago uses its selection device on a cutoff and rank-order basis although it has evidence that individuals can perform the job well who do not score at the top of the rank-order list.[196] The city hired its one female firefighter

---

[188] *Id.* at 758 n.8.
[189] *Id.* at 756 n.5.
[190] 536 F. Supp. at 210.
[191] Guardians Ass'n v. Civil Serv. Comm'n, 630 F.2d 79, 100 (2d Cir. 1980).
[192] 536 F. Supp. at 210.
[193] *Id.* at 211.
[194] *Id.* at 211-12.
[195] *Id.* at 212.
[196] *See* FINAL REPORT, *supra* note 139.

when it went far down its eligibility list during a firefighter strike.[197] Although she and many of the men hired at that time scored just above the cutoff score, they have performed competently.[198] Similarly, although the City of Buffalo uses its physical performance examination on a rank-order basis, the five women it hired as the result of a court order have performed competently even though they scored just above the cutoff score.[199]

In order to assess whether a test is valid on a rank-order basis it is crucial to include women in the job analysis. To date, no job analysis has included female participants. Existing studies have shown that firefighters have the same or worse physical conditioning than the average male.[200] On the basis of such findings, jurisdictions have set cutoff scores above that of the average man.[201] No jurisdiction has ever inquired whether the average woman's physical ability is also sufficient to be a firefighter.

The Uniform Guidelines require that validity studies include a sample of the protected class upon whom the test has had adverse impact, whenever practical.[202] The practicality exception recognizes that studies of job incumbents cannot include groups historically excluded from the job. Despite this requirement, test developers have not included women in their samples even when they were studying a group of potential job applicants rather than job incumbents.[203]

---

[197] *Berkman*, 536 F. Supp. at 211 n.22.

[198] *Id.*

[199] *See* testimony of Margaret Keane (Buffalo firefighter), United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y. hearing held Mar. 22, 1983).

[200] The *Berkman* court noted that physical testing results showed that "on a variety of measures of maximum physical capacity firemen rate no better than the average American male." *Berkman*, 536 F. Supp. at 212.

[201] *See, e.g., id.* at 198 (noting that many firefighters would have failed exam; no women used in the sample).

[202] Uniform Guidelines, *supra* note 15, § 1607.14B(4).

[203] For instance, the City of Chicago used three pilot studies in developing its test, with women only included in the final study. FINAL REPORT, *supra* note 139, at iii-vii. Because no women were in the first two pilot studies and no racial minorities other than blacks were in the second pilot study, the studies failed to control sufficiently for body size and composition. The flexed-arm hang, for example, was chosen as a measure of percentage of body fat because of the concern for obesity control among firefighters. *Id.* at 58. Because women, on average, have a lower center of gravity than men, the flexed-arm hang is a poor measure of obesity in women. Moreover, because women have a biologically higher percentage of body fat than men, W. MCARDLE, *supra* note 16, at 389, it is inappropriate to use the same standard of percentage of body fat for women and men to measure obesity. Women and men with the same percentage body fat are not relatively the same in terms of obesity. This problem would

For instance, the City of Seattle failed to explore women's physical abilities sufficiently in setting cutoff scores.[204] This is especially disappointing because Seattle made an attempt to include women in its study. The city had developed physical performance goals based on a study of three female job applicants and two classes of female prerecruit firefighters.[205] The city undertook a study to validate those goals by evaluating the physical capabilities of firefighters as they proceeded from the city's training program to positions as career firefighters. Thirteen women started in the first class and eleven finished.[206] They had attained ninety-one to ninety-seven percent of the established strength goals and were adjudged marginally strong enough.[207] Nevertheless, "lack of sufficient strength" induced ten of the eleven to resign from positions as career firefighters, according to the report.[208] The other recruit resigned due to an injury. Seven women were hired for class two. None of them met the goals, although the study found them

---

be particularly significant for black women because, on average, they are twenty pounds heavier than white women between the ages of 35 and 65. Abraham, *Height and Weight of Adults 18-74 Years of Age in the United States*, in 3 ADVANCE DATA FROM VITAL AND HEALTH STATISTICS (Nov. 19, 1976). If that additional weight is located primarily in the lower half of the body, then black women would be at a particular disadvantage in the flexed-arm hang. Because the first pilot study had revealed that black men had significantly more fat-free body weight than Spanish surnamed individuals, it was important to include Spanish surnamed individuals, as well as women, in each of the pilot studies.

The third stage of the Chicago validity study confirmed that the test was not equally predictive for all individuals. The subjects in this stage of the study included 9809 firefighter applicants, of whom 9763 were male (white = 6255, black = 2946, Spanish surname = 483, American Indian = 23, Asian American = 16, and Other = 40), 35 were female and 11 did not indicate their sex. FINAL REPORT, *supra* note 139, at 80. The test implementation data revealed that Spanish surnamed individuals did considerably worse than the other racial groups. *Id.* No explanation was given for that disparity nor were any further studies conducted. The results also indicated disparate impact on the basis of sex. *Id.* Rather than explore why the differences occurred, the test developers stated: "Men typically have greater muscle strength and cardiovascular endurance than women. . . .The extent to which these differences are culturally related is not known." *Id.* at 89. No attempt was made to determine if the test was equally predictive of men's and women's abilities or if the test measured not strength but rather skills which could be easily taught.

[204] *See* T. Doolittle, Validation of Physical Requirements for Firefighters, prepared for the Seattle Fire Department & Office of Management and Budget (Mar. 14, 1979) [hereafter Seattle Study].

[205] *Id.* at 7.

[206] *Id.* at 7-8.

[207] *Id.* at 8.

[208] *Id.* at 8.

"quite strong relative to their lean body weight."[209] The study predicted that only two might successfully complete firefighter recruit training. It observed that "as a group they are experiencing difficulties due to insufficient upper body strength."[210] It was assumed that all difficulties were attributable to upper body strength rather than skill, even though these women purportedly had considerable upper body strength.

The city then studied a group of incumbents to establish norms.[211] It chose ninety subjects from individuals who were not on "disability" and who were below the rank of captain.[212] The study did not consider whether the individuals eliminated were competent firefighters despite their limitations.[213] Additionally, it established norms above the scores of many incumbents to account for aging.

A major problem with the study's conclusion was the failure to consider that strength may not have been the determinant factor in the difficulties the women allegedly experienced. In fact, the women had as much strength as many of the successful male firefighters, not including the men dropped from the analysis. Moreover, the study assumed that the aging factor would affect men and women equally. Because many women have experienced physical activity for the first time quite recently and are voluntarily engaging in such activity in record numbers despite increasing age, those assumptions may not be equally valid for men and women. Finally, the study did not consider the possibility that

---

[209] *Id.* at 9.

[210] *Id.* at 10.

[211] *Id.*

[213] *Id.* The study eliminated six subjects because of orthopedic problems. *Id.* A seventh individual was later disqualified because he was taking cortisone. *Id.* at 10-11. Additional subjects were chosen and a final sample of 88 was then studied. *Id.* at 11.

[213] On the basis of this data, the study concluded that the original goals for upright row, forward rise, lateral rise, pull-ups, and aerobic power were higher than the scores of the average male incumbent. *Id.* at 15. The study also found that the original goals for military press, handgrip, and overall sum were higher than 10% of the incumbents' scores. *Id.* Only the values for bicep curls and squats were equivalent to the values for the lowest scoring incumbents. *Id.* Despite these findings, the test validaters recommended increasing the goals for bicep curls, squats, and military press, leaving handgrip the same and lowering upright row, based on considerations of the effect of aging. *Id.* at 15-17, 18-23. The goal for pull-ups was also lowered because the study found that there was no significant aging effect with pull-ups. *Id.* at 23. Finally, the aerobic power goal was not modified despite a finding that the mean for the incumbents was much lower than the benchmark and only equal to the mean aerobic power for men of 35.4 years of age. *Id.* at 15, 24-25. No consideration was given to the fact that the men's average strength may also have been average for men in their age group — disputing that firefighters have more than average strength or endurance.

increases in skill developed with age might counterbalance any losses in strength or speed. By stereotyping the nature of the job's physical demands and women's limitations, the test validaters preordained their result that higher physical requirements were necessary. In fact, they had no evidence of job performance by the women who they asserted were too weak. The study's only data consisted of projections of women's failure based on limited observations of women who were quite strong. Further, the test validaters had no evidence that male firefighters, who the study presumed were getting weaker with age, were not performing competently.[214]

Studies need to examine the growing number of successful female firefighters to determine what techniques, physical attributes, and cognitive abilities they use to be successful firefighters. Rather than begin with the assumption that physical strength cutoffs must be set above those of the strongest women in the applicant pool, studies must begin with the assumption that successful female firefighters do exist.

IV. Possibilities for the Future: Consideration of Alternate Selection Devices

Many jurisdictions have begun to hire police officers without regard to race or sex, often under court-imposed goals and timetables.[215] Al-

---

[214] By contrast, the *Berkman* court based its analysis on the assumption that older firefighters, who might be getting weaker during their careers, are nevertheless competent. It then used that fact to downplay performance at the limits of strength:

> [F]irefighting takes its toll, not as a result of failures of maximum strength or speed, even at critical moments, but rather through the physical demands extending over long periods of time which necessitate paced performance at less-than-maximum levels. This explanation makes sense of a number of otherwise puzzling features of firefighting, namely, the ability of firemen to continue to do their work competently over an entire working career and the results of physical testing of incumbent male firemen which show that on a variety of measures of maximum physical capacity firemen rate no better than the average American male. It also makes sense in terms of the recognized dangers of firefighting and their unpredictability which, as numerous witnesses testified, make hazardous in the extreme performance at top speed or at the limits of strength or capacity. Not only does maximal performance comprise the firefighter's ability to pace performance over the long periods of time during which physical demands are being made on the body. In addition, the very unpredictability of fire and the instability of burning structures call upon qualities of foresight, endurance, and pacing not examined by tests of maximum physical strength.

Berkman, 536 F. Supp. at 212.

[215] *See, e.g.*, Harless v. Duck, 619 F.2d 611 (6th Cir. 1980); Blake v. City of Los

though these police departments may initially have been reluctant to take that step, studies indicate that the introduction of women and minorities has improved the quality of law enforcement.[216] Nevertheless, these jurisdictions generally have not developed selection criteria that function without disparate impact upon a protected class, because they have been able to use goals and quotas to achieve no disparate impact. By contrast, in firefighting, women have made few inroads in employment opportunities — few fire departments are under court order to hire women and few departments have tried to devise selection devices that would not have an impact upon women.

It is important to consider what nondiscriminatory selection devices might be used in police or firefighter hiring without goals and quotas. The United States Justice Department is seeking the elimination of goals and quotas from consent decrees with state and local governments.[217] Unless nondiscriminatory selection procedures are developed, many departments may revert to using selection devices that disparately impact women or minorities. In addition, even if an employer can validate a selection device for a particular use, the Uniform Guidelines provide that a plaintiff from a protected class may show that other procedures are available that would produce a smaller disparate impact on the protected class.[218] Hence, plaintiffs often have an opportunity to present an alternative nondiscriminatory hiring procedure.

Alternatives to rank-order physical abilities or written examinations do exist. Written and physical abilities examinations could be scored on a cutoff basis set at the job's minimum physical and cognitive requirements. If a training program that teaches the skills necessary to pass the physical and written examinations preceded the administration of

---

Angeles, 595 F.2d 1367 (9th Cir. 1979).

[216] *See, e.g.*, Dothard v. Rawlinson, 433 U.S. 321, 344 (1977) (Marshall, J., dissenting) (observing that the presence of female guards did not lead to a single incident amounting to a serious breach of security in any Alabama institution); *Harless*, 619 F.2d at 616 (noting that physical ability test was deleted without any detrimental effect); *see also* N.Y. Times, Oct. 13, 1985, at E6, col.3 (describing successful introduction of women into police force and sanitation department).

[217] *See, e.g.*, Equal Employment Opportunity Comm'n v. Local 638, 36 Fair Empl. Prac. Cas. (BNA) 1466 (2d Cir. 1985); *In re* Birmingham Reverse Discrimination Employment Litig., 37 Fair Empl. Prac. Cas. (BNA) 1 (N.D. Ala. 1985). However, these efforts may not be successful. *See* Wygant v. Jackson Bd. of Educ., 54 U.S.L.W. 1177 (May 20, 1986) (concurring opinion validates affirmative action plans with legitimate and remedial purposes).

[218] Albemarle Paper Co. v. Moody, 422 U.S. 405, 425 (1975); *see also Harless*, 619 F.2d at 617; Officers for Justice v. Civil Serv. Comm'n, 395 F. Supp. 378, 384 (N.D. Cal. 1975) (requiring the consideration of less restrictive alternatives).

the examinations, the examinations would have minimal disparate impact on women and minorities. Women would disparately benefit from a brief physical training program that stressed the skills needed to pass physical examinations. Similarly, a training program for a written examination would disparately benefit racial minorities who often have not been taught the necessary test-taking techniques.[219] This is the type of relief ordered in *Berkman*.[220] The cities of San Diego and Los Angeles have voluntarily undertaken such training programs in hiring firefighters.[221]

A physical training program could greatly reduce the adverse impact of a physical abilities selection device on women because women, on average, benefit more from physical training than men.[222] Women benefit greatly from training because most tests have a significant skill component which can readily be taught. As the *Berkman* court observed: "tricks of the trade that can be taught do not represent appropriate abilities to measure to determine whether the applicant can perform successfully on the job."[223] Applying that analysis, the court found that the body carry, stair climb, obstacle course, and hose coupling test

---

[219] *See generally* S. GOULD, THE MISMEASURE OF MAN (1981)

[220] *Berkman*, 536 F. Supp. at 218.

[221] For example, since 1980, the selection procedure for San Diego has included a written test and a physical abilities test. A study guide was given to all applicants who signed up for the written examination. It contained material to be studied and memorized for the examination and sample questions for each test area on the examination. The physical abilities test included tests of endurance, strength, and flexibility. The test was administered on a pass-fail basis. A physical agility preparation class was offered several months before the examination. Based on the examination results, applicants were divided into three approximately equal categories: 90 and above, 80 through 89, and 70 through 79. *See* Report on the San Diego Fire Fighter Physical Ability Test, in B. Lopez, Recruitment Publicity Plan (Oct. 9, 1984) (copy on file with *U.C. Davis Law Review*).

Although the author approves of the general technique used by San Diego to select firefighters, she does not necessarily approve of the specific items used on the written or physical abilities examinations. More information would be needed to make that evaluation.

See also N.Y. Times, Oct. 13, 1985, at E6, col. 3 (describing pass-fail written and physical examinations used for selecting sanitation workers in conjunction with a lottery).

[222] *See* ARMY REPORT, *supra* note 93, at 10 (reporting that the physical performance discrepancies between men and women "were reduced following training since women usually start training at lower levels of their physical potential and improve more in response to a training program"); *see also* N.Y. Times, Sept. 9, 1982, at B2, col. 5 (reporting that 16 of 24 women were able to pass the New York City firefighters physical performance examination after undergoing training).

[223] 536 F. Supp. at 208.

items in the fire department's physical performance examination involved skills that could readily be taught.[224]

One item in the New York City obstacle course was an eight-foot wall.[225] For all women but one taking the test, that item was an absolute bar to completing the exam successfully.[226] The one woman who did get over the wall testified she used a simple technique she had been taught in a private training school. The technique did not involve any of the abilities purportedly measured.[227] Because of the socially created differences between men and women's physical opportunities,[228] it is more likely that men have been taught the skills necessary to complete a physical performance examination.[229]

Physical training could also help women learn the most efficient way to make use of body strength. In San Diego, for instance, the fire department has observed that women have greater success in carrying ladders and moving other heavy objects if they take advantage of their lower body strength.[230] The city's technique of using lower body strength is actually safer for all individuals, male and female, because it puts less stress on the lower back. Hence, more exploration of techniques relying less on upper body strength may benefit the entire workforce.

If a rank-order selection device were needed to select among the candidates who possess the ability to perform the job's physical and cognitive requirements, other nondiscriminatory selection devices could be used to rank-order candidates for final selection. Relying on selection techniques such as psychological examinations or interviews would permit employers to assess whether job applicants possess the psychological

---

[224] *Id.* at 207-08.

[225] *Id.* at 194.

[226] *Id.* at 194 n.12.

[227] *Id.*

[228] *See supra* text accompanying notes 27-93.

[229] For example, at the hearing in United States v. City of Buffalo on the validity of the city's physical performance examination, one of the experts testified that several of the women hired by New York City were having trouble starting the motor on the "jaws of life," a tool used to open steel doors, such as those on automobiles. The city claimed that the women were not strong enough. The expert asked for an opportunity to train them on the technique needed to start the machine. In a brief afternoon session he successfully taught them the required technique. Their problem was one of skill, not strength. *See* Trial Transcript of Testimony of W. McArdle, United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y. Mar. 22, 1983).

[230] Conversations of Ruth Colker with United States Department of Justice attorneys who visited the San Diego Fire Department (Jan. 1984).

traits necessary to perform under stressful conditions.[231] Although interviews historically have been used to exclude women and minorities because of their subjective structure, interviews are not inherently flawed as a selection device. Highly structured interviews can minimize interviewer bias and result in the selection of highly qualified candidates.[232] The use of these selection techniques might not only lessen disparate impact upon women and minorities in hiring, but also result in an improved overall workforce through selection of individuals with the physical, cognitive, and psychological traits necessary to be police officers or firefighters.

CONCLUSION

Historically, women have been absolutely excluded from physical activity and "physically demanding" professions. Women's exclusion from these activities has helped perpetuate the stereotypes that only males possess the highly valued physical attributes of strength and speed. As women's entry into the world of athletics and police work has begun to challenge conceptions of men's and women's proper roles, the doors to the firehouse have remained closed to women. The development of an exclusionary device more subtle than absolute exclusion — rank-order use of physical performance examinations — is the cause of this. Ironically, the Department of Justice, which has been on the forefront of opening up the police profession to women by demonstrating the unlawfulness of such examinations, is at the firehouse door, barring entry. Nevertheless, extensive precedent regarding invalid written examinations should be used to strike down invalid, exclusionary physical abilities selection devices, bringing women into yet one more traditionally male profession.

---

[231] *See generally* Trial Transcript of Testimony of D. Gebhardt, United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y. Mar. 23, 1983). Another possibility, which has been successfully attempted in San Diego, is to administer the physical examination on a pass-fail basis and to divide scores on the written examination into three categories. Applicants were then selected from the top-third of the test takers, but not necessarily on a rank-order basis within that group. *See supra* note 221.

[232] Trial Transcript of Testimony of D. Gebhardt, United States v. City of Buffalo, No. 74 Civ. 195 (W.D.N.Y. Mar. 23, 1983).