Suppose you supervise the police department of a large American city. In response to rising crime, the city council has authorized you to hire more officers. Within days your office is flooded with applications for just a handful of positions. How can you possibly choose from among them?

One possibility is to screen those applicants initially to eliminate those who seem unlikely to become successful candidates. Imposing minimum education and literacy requirements might remove some applicants from consideration. Another option would be to screen based on physical ability. After all, many police officers must apprehend fleeing criminals, and skills such as strength and stamina could be measured by requiring applicants to drag or lift a certain amount of weight or complete a timed run. However, such a test might run afoul of employment discrimination law, specifically Title VII. Title VII protects minorities and women from employment practices that discriminate against them. Since men tend to outperform women on most physical tests, more men than women will be hired as a result of using such tests. Consequently, tests based on physical strength can create an unlawful disparate impact.

The Civil Rights Act of 1991 provides for a “business necessity” or “job related” defense to disparate impact claims under Title VII. Many employers rely on this defense when contesting their liability. The Act allows an employer to prevail if it proves “that the challenged practice is job related for the position in question and consistent with business necessity.” However, the exact contours of this defense are still unclear, and lower courts have used a number of different standards when evaluating the defense, thereby confusing plaintiffs and defendants alike.

Part I of this Comment reviews the history of disparate impact analysis and the business necessity defense under Title VII and the Civil Rights Act of 1991. Part II examines one area where the business necessity defense crops up quite frequently, the use of physical ability

3 Id.
tests, which tend to favor men over women. Part III notes at least four different standards that have been applied by the courts in the physical ability test context. Part IV proposes a two-prong test that best reflects the command from Congress that any test with a disparate impact be both "job related" and "consistent with business necessity." The Comment then applies the proposed test to the physical ability test context and to disparate impact litigation in general.

I. BACKGROUND

A. History of Title VII

Title VII of the Civil Rights Act of 1964 prohibits job discrimination on the basis of race, sex, religion, or national origin. When originally enacted, it was unclear whether Title VII would treat as an "unlawful employment practice" a job requirement or employment test that, while facially neutral, tended to disproportionately exclude women or racial minorities from a particular position. The Supreme Court answered that question in the seminal case of Griggs v Duke Power Co, where the Court concluded that even with no finding of discriminatory intent, an employment test that had a disparate impact on a protected class could violate Title VII. In Griggs, the Court held invalid a requirement that power plant workers who wished to be promoted to coal handling positions either have a high school diploma or pass a standardized intelligence test. The Court found that such a requirement tended to hinder the opportunity for African-Americans to earn promotions, and it based its decision, in part, on an analysis of the workforce of the power plant in question, which showed a wide disparity between the number of white and African-American coal handlers. The Court was careful to note that not all tests causing a disparate impact in employment opportunities would violate Title VII.

4 Id.
6 The legislative history indicates that most members assumed Title VII would not apply to such disparate impact cases. See Andrew C. Spiropoulos, Defining the Business Necessity Defense to the Disparate Impact Cause of Action: Finding the Golden Mean, 74 NC L Rev 1479, 1481 (1996) (noting that the Court may have intended only to hold employers liable for "intentional discrimination"). See also 110 Cong Rec S 7213 (Apr 8, 1964) (interpretive memorandum of Senators Clarke and Case) (declaring that employers need not abandon "bona fide qualification tests" if members of some groups outperformed members of other groups).
8 See id at 432 (concluding that "good intent or absence of discriminating intent does not redeem employment procedures or testing mechanisms that ... are unrelated to measuring job capacity"). Griggs is a statutory rather than a constitutional decision. The Equal Protection Clause of the Fourteenth Amendment is violated only by purposeful discrimination, not by actions that merely create a disparate impact. Washington v Davis, 426 US 229, 239 (1976).
10 Id at 426, 430 n 6.
11
However, the Act had "placed on the employer the burden of showing that any given requirement must have a manifest relationship to the employment in question." In a key passage the Court explained, "The touchstone is business necessity. If an employment practice which operates to exclude Negroes cannot be shown to be related to job performance, the practice is prohibited." Thus in post-Griggs disparate impact cases, one important issue is determining exactly which tests that create a disparate impact are valid and which violate Title VII. The phrases "business necessity" and "job related" have become critical to this analysis, despite the apparent redundancy of these two terms. After all, in common parlance, if an employment practice is a business necessity, it seemingly must also be job related.

In its subsequent jurisprudence, the Supreme Court clearly held that not all tests creating a disparate impact were necessarily a violation of Title VII. Nevertheless, the Court did little to clarify exactly when an employment test that had a disparate impact was nonetheless a reasonable measure of job performance. The Supreme Court next addressed the issue in Albemarle Paper Co v Moody, where it seemingly adopted a strict standard that would likely render most physical and written tests invalid. In Albemarle, a paper plant sought to use in making its hiring and promotion decisions a written skills test on which white workers outperformed racial minorities. In affirming a lower court decision striking down this test, the Court established the burden-shifting analysis that has become critical to all Title VII disparate impact claims. First, the plaintiff must establish that the tests in question select applicants for hire in a "pattern significantly different from that of the pool of applicants." If the plaintiff makes this showing of disparate impact, the burden shifts to the employer to validate that its test is "significantly correlated with important elements of

---

11 Id at 432.
12 Id at 431.
13 422 US 405 (1975).
14 Many written tests tend to discriminate against African-Americans or Hispanics, who on average are less educated than whites. See NAACP v New York, 413 US 345, 363 (1973) (finding that literacy tests have a greater impact on blacks because of higher white literacy rates); John A. Powell and Marguerite L. Spencer, Remaking the Urban University for the Urban Student: Talking About Race, 30 Conn L Rev 1247, 1287-88 (1998) (noting that black and Hispanic scores on the SAT and LSAT lag behind white scores). Similarly, physical tests that measure skills such as strength and stamina tend to discriminate against women who on average are weaker and slower than men. See Evans v City of Evanston, 881 F2d 382, 384 (7th Cir 1989) ("Since men are on average stronger and faster than women, the higher the passing score on a... physical agility test... the smaller the percentage of women likely to pass it."); Timothy H. Goldsmith and Owen D. Jones, Evolutionary Biology and Behavior: A Brief Overview and Some Important Concepts, 39 Jurimetrics J 131, 134 (1998) (finding that men are stronger and 8 percent larger than women).
15 Albemarle, 422 US at 410-11.
16 Id at 425.
work behavior which comprise or are relevant to the job." Since the plaintiffs in *Albemarle* had demonstrated a disparate impact in hiring and the paper plant was unable to point to any correlation between test scores and job success among those who had passed the test, the Court determined that the test was invalid.

In *Dothard v. Rawlinson*, the Supreme Court for the first (and to date only) time considered the legality of an employment test that had a disparate impact on women. This case also seemed to espouse a strict standard, although the evidence there was more conflicting. The plaintiff in *Dothard* failed to meet a requirement that all Alabama prison guards be at least five-foot two inches and weigh at least 120 pounds. Such a rule would have screened out a substantial percentage of the female population, but very few men. The state asserted that its rule was justified because prison guards needed to be strong, and taller, heavier people tended to be stronger. However, the defendants failed to introduce any statistical evidence, essentially urging the adoption of a rational basis standard for employer defenses in the disparate impact context.

The Court declined to embrace such a standard, concluding that once the plaintiff established a prima facie case of disparate impact, the burden shifted to the defendant to produce statistical evidence justifying the need for the test or standard. In a footnote frequently cited by those attacking physical tests, the Court implied that employers could use discriminatory testing mechanisms only under very limited circumstances: "[A] discriminatory employment practice must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge." While alluding to this strict standard, the Court also stated that if strength really were a bona fide "job-related quality," it should not be difficult for the state to validate a test that measured strength directly instead of indirectly through reliance on height and weight measurements.

---

1 Id at 431, quoting 29 CFR § 1607.4(c).
2 *Albemarle*, 422 US at 431–32. The EEOC has issued guidelines detailing the evidence an employer must present for a test to be deemed "validated." See 29 CFR § 1607.5(B) (1999).
4 Id at 331 (finding that height and weight requirements had a disparate impact).
5 Id at 323–24.
6 Id at 329 (noting that the height and weight requirements would exclude 33.29 percent of all women between 18–79 years of age, but only 1.28 percent of all men).
7 Id at 331.
8 Id at 331 n 14.
9 Id. See also id at 339–40 (Rehnquist concurring) (discussing appellants' near total lack of evidence).
10 Id at 331 (“If the employer discerns fallacies or deficiencies in the data offered by the plaintiff, he is free to adduce countervailing evidence of his own.”).
While implying a strict standard in disparate impact cases, the Supreme Court did uphold two employment requirements that were challenged by African-Americans in the late 1970s. In New York City Transit Authority v Beazer, the Court determined that there was a "manifest relationship" between a rule against hiring methadone users and the need to hire employees who would be "safety sensitive." Surprisingly, in light of Dothard, the Court did not assess whether the ban on methadone users was necessary for safe job performance. Indeed, Justice Stevens' majority opinion found that the transit authority's goals of safety and efficiency were "significantly served by—even if they do not require—[the methadone] rule."

In Washington v Davis, the Supreme Court upheld a literacy test for police trainees in the face of challenges under the Equal Protection Clause and a District of Columbia anti-discrimination statute. The Court first concluded that a disparate impact cause of action could not be brought under the Equal Protection Clause, which requires a finding of discriminatory intent. It then went on to affirm the district court's finding in favor of the employer on the statutory claim. In upholding the literacy test, the Court noted that "some minimum verbal and communication skill would be very useful, if not essential, to satisfactory progress" on the job. While the Court's interpretation of the statute at issue is obviously not controlling on Title VII analysis, the Court did assume that the relevant standards were "similar to those obtaining under Title VII" and not "foreclosed by either Griggs or Albemarle."

B. Wards Cove and the Civil Rights Act of 1991

Amidst this muddled precedent, lower courts faced with employment screening practices that created a disparate impact on a protected class adopted a variety of formulations of the "business necessity" or "job related" defense. To resolve these inconsistent interpre-

30 Id at 587 n 31, quoting Griggs, 401 US at 432.
31 440 US at 587 n 31.
33 At the time of the suit, Title VII did not apply to the federal government. See Washington, 426 US at 236 n 6.
34 Id at 238-39.
35 Id at 248.
36 Id at 250. By contrast, there was no demonstration in Griggs of any need for verbal skills in the job of coal handling. See Griggs, 401 US at 433.
38 Id at 251.
39 See Part III.
tations, the Supreme Court attempted to retreat from the strict business necessity test enunciated in some of its earlier opinions by setting out a clear pro-employer standard in *Wards Cove Packing Co v Atonio.* In *Wards Cove,* minority cannery workers alleged that they were generally given lower paying jobs and restricted in their chances for promotion because of the disparate impact created by the defendant company's hiring practices. One of the questions presented was, assuming the plaintiff made out a case of disparate impact, which party had the "burden to prove that any disparate impact caused by [the defendant's] hiring and employment practices was justified by business necessity." The lower court had found that the burden of proving that a discriminatory test was justified should be placed on the defendant. The Supreme Court, while acknowledging that its precedent could be read either way on the matter, reversed, holding that the burden of proof should be placed on the plaintiff alleging the discrimination, and not the employer.

The primary holding of *Wards Cove* was this shifting of the burden of proof on the business necessity defense. But the Court attempted to do more than merely shift the procedural burden of proof; it also sought to lay out the substantive elements of the defense itself. Once a plaintiff had demonstrated that an employment practice created a disparate impact, "the dispositive issue is whether a challenged practice serves, in a significant way, the legitimate employment goals of the employer."

"This standard itself could be seen as vastly more pro-employer than the *Dothard* requirement that a challenged practice be necessary for safe and efficient job performance, and more in line with the deferential standard of *Beazer.*"

Dissatisfied with *Wards Cove* and numerous other Supreme Court decisions from the October 1988 Term cutting back on the protections of Title VII, Congress was spurred into action. It passed the

---

41 Id at 647–48.
42 Id at 649.
43 Id at 648.
44 Id at 660 ("We acknowledge that some of our earlier decisions can be read as suggesting otherwise.").
45 Id at 659.
46 Id.
Civil Rights Act of 1991\(^{48}\) to statutorily reverse these decisions. The Act attempted to codify Supreme Court precedent prior to *Wards Cove* by clarifying that if a plaintiff demonstrates disparate impact caused by an employment practice, he or she will prevail unless the employer proves “that the challenged practice is job related for the position in question and consistent with business necessity.”\(^{49}\) Even then, the plaintiff can preserve his or her case by demonstrating “an alternative employment practice” that would not have such a disparate impact.\(^{50}\)

While it is clear that the Act shifted the burden of proof on the business necessity defense to the defendant, it is less clear if the Court’s substantive interpretation of the defense itself, as set out in *Wards Cove*, was also overruled. The use of legislative history to resolve this question is foreclosed since the Act explicitly provides that legislative history may not be considered in interpreting its provisions aside from a single memorandum.\(^{51}\) The relevant memorandum states only that “[t]he terms ‘business necessity’ and ‘job related’ are intended to reflect the concepts enunciated by the Supreme Court in *Griggs v. Duke Power Co.*, and in the other Supreme Court decisions prior to *Wards Cove Packing Co. v. Antonio.*”\(^{52}\) However, because of the confused state of the cases between *Griggs* and *Wards Cove*, the exact parameters of the business necessity defense are still far from resolved.\(^{53}\)

II. PHYSICAL ABILITY TESTS

The lack of clarity in both Supreme Court jurisprudence and the Civil Rights Act of 1991 on the contours of the business necessity defense has led to widely varying applications in the lower courts. One clear example of this problem appears in the realm of physical ability


\(^{50}\) 42 USC § 2000e-2(k)(1)(A)(ii).


\(^{52}\) 137 Cong Rec S 15276 (Oct 25, 1991).

\(^{53}\) See *Allen v Entergy Corp*, 181 F3d 902, 904 (8th Cir 1999) (finding that there must be a manifest relationship between the employment practice and the tested skill); *Lanning v South-eastern Pennsylvania Transportation Authority*, 181 F3d 478, 481 (3d Cir 1999) (requiring employer to prove test measures the “minimum qualifications necessary for successful [job] performance”), cert denied, 120 S Ct 970 (2000); *NAACP v Town of East Haven*, 70 F3d 219, 225 (2d Cir 1995) (requiring manifest relationship between employment practice and test).
tests. Physical tests such as obstacle courses, \(^{54}\) mile runs, \(^{55}\) wall climbs, \(^{56}\) or hose drags \(^{57}\) have been and often still are used as screening mechanisms for occupations such as police officer and firefighter. \(^{58}\) The validity of such tests has been subject to much litigation since the Griggs decision and the Equal Employment Opportunity Act of 1972. \(^{59}\) There are at least two reasons why these tests are frequently litigated. First, because men have greater exposure to athletic opportunities and are generally quicker and stronger than women, most physical tests have the effect of preferring men over women, creating an inherently disparate impact. \(^{60}\) Because of this, the central issue in most physical test litigation is not the existence of a disparate impact, but whether or not the test is job related and consistent with business necessity. In addition, historically male-only police and fire department positions were only reluctantly opened to women and still suffer from woeful female representation. \(^{61}\) Consequently, it is often feared that such tests were put into place solely to continue discriminatory hiring practices through facially neutral criteria. Thus at times tests may be used to conceal disparate treatment that is actually occurring—an illegal desire by some police or fire departments to keep out women. \(^{62}\) On the other hand, physical tests do provide a neutral screening mechanism.

---

54 See Harless v Duck, 619 F2d 611, 616 (6th Cir 1980).
55 See Lanning v Southeastern Pennsylvania Transportation Authority, 181 F3d 478, 482 (3d Cir 1999).
57 See Pietras v Board of Fire Commissioners, 180 F3d 468, 471 (2d Cir 1999).
58 Physical tests can of course be used in the private sector, but they appear to be quite rare. For one example, see EEOC v Simpson Timber Co, 1992 US Dist LEXIS 5829, *8–11, *13–14 (W D Wash) (concluding that although a physical test for private timber company applicants had a legitimate business purpose, alternatives with a less disparate impact existed).
59 Pub L No 92-261, 86 Stat 103, 103, codified as amended at 42 USC §§ 2000e et seq (1994) (extending Title VII to cover state and local governments, which frequently use physical tests in making hiring decisions).
60 See Zamen v City of Cleveland, 906 F2d 209, 211 (6th Cir 1990) (stating that men tend to excel at anaerobic functions such as speed and strength). See also note 14 and Ruth Colker, Rank-Order Physical Abilities Selection for Traditionally Male Occupations as Gender-Based Employment Discrimination, 19 UC Davis L Rev 761, 767–72 (1986) (describing how prior to Title VII, women had limited exposure to athletic events and were discouraged from participating in highly demanding physical sports).
61 See Tim Frenzler, Rebuilding the Walls: The Impact of Police Pre-Entry Physical Ability Tests on Female Applicants, 7 Current Issues Crim Just 314, 324 (1996); Colker, 19 UC Davis L Rev at 773–75 (cited in note 60) (discussing the resistance to admitting women to the Los Angeles Police Department); Kevin Flynn, Despite Recruiting, Few Women Do Well in Firefighting Tests, NY Times A1 (Feb 3, 2000) (noting that only thirty-six of New York’s 11,000 firefighters are female).
62 See Legault v Zambarano, 105 F3d 24, 26 (1st Cir 1997) (upholding Rule 11 sanctions against town officials and their attorney who had used a physical test as a sham to conceal that “[h]iring decisions [for fire department positions] were, in fact, based on undisclosed, subjective criteria within the discretion of the defendants”). See also Hurley v Atlantic City Police Department, 174 F3d 95, 104 n 5 (3d Cir 1999) (noting frequency of sexual harassment of female police officers by male coworkers).
for government agencies forced to decide between hundreds of applicants in cases where the use of subjective criteria or educational and literacy requirements often lead to disparate impact challenges from racial minorities.

At least twelve published cases have appeared over the last twenty years evaluating claims asserted by females seeking to invalidate physical ability tests. These decisions have diverged widely, splitting almost in half on findings of liability and delineating numerous standards for applying the business necessity defense. The lack of clarity in the area makes it harder both for employers to design valid tests and for applicants to successfully press claims of genuine gender discrimination. The Civil Rights Act of 1991 rejected the clear pro-employer stance of Wards Cove, but, in returning the law to that of Griggs and its progeny, Congress left a muddled standard that courts have struggled to apply.

III. STANDARDS FOR ANALYZING PHYSICAL ABILITY TESTS UNDER TITLE VII

Because the Supreme Court precedent is so confusing and the terms “business necessity” and “job related” are seemingly redundant, lower courts have struggled to articulate a rule for analyzing physical test cases. Essentially, four major standards have developed. From most to least deferential to employers they are: manifest relationship, the Spurlock public safety doctrine, close approximation, and minimum qualifications. Because these approaches include standards that show great deference to employer tests and those that would ban virtually any physical test used as a prerequisite for employment, they can yield substantially different results when applied to similar physical tests.

A. Manifest Relationship

The most lenient standard merely requires the defendant employer to prove a manifest relationship between the physical examination and successful job performance. This standard is based on the

63 Compare Lanning, 181 F3d 478; Pietras, 180 F3d 468; Berkman v City of New York, 705 F2d 584 (2d Cir 1983); Harless, 619 F2d 611; Blake v City of Los Angeles, 595 F2d 1367 (9th Cir 1979); Legault v aRusso, 842 F Supp 1479 (D NH 1994); Simpson Timber Co, 1992 US Dist LEXIS at 5829; Brunet v City of Columbus, 642 F Supp 1214 (S D Ohio 1986) (invalidating physical ability tests), with Zamlen, 906 F2d at 209; Berkman v City of New York, 812 F2d 52 (2d Cir 1987); Eison v City of Knoxville, 570 F Supp 11 (E D Tenn 1983); Hardy, 576 F2d 1342 (upholding physical ability tests). In theory, any protected class could sue for the disparate impact of a physical ability test, but all published cases have involved claims brought by female applicants asserting gender bias.

64 See Harless v Duck, 619 F2d 611, 616 n 6 (6th Cir 1980) (discussing the “manifest rela-
contention in \textit{Griggs} that an employer must show "that any given re-

quirement [has] a manifest relationship to the employment in ques-
tion."\textsuperscript{65} Thus, if the employer demonstrates a legitimate business pur-
pose and shows that the test implemented will effectively carry out that business purpose, the test will survive challenge.\textsuperscript{66} The manifest relationship standard was used in \textit{Eison v City of Knoxville},\textsuperscript{67} where a female police trainee challenged a requirement that all recruits be able to pass a battery of tests including "sit-ups, push-ups, leg lifts, squat thrusts, pull-ups, and a two mile run."\textsuperscript{60} While the court did re-

quire the defendant police department to establish that the test was "job related,"\textsuperscript{69} this meant only that the score used as a cutoff was not arbitrary but seemed in some way rational, not that there was any spe-
cific correlation between test success and job success.\textsuperscript{70} Thus, in \textit{Eison}, the court found that a police department could rationally conclude that attributes such as upper and lower body strength and endurance were important to the job, and that these attributes could be measured through the battery of tests established.\textsuperscript{71} This rendered the test valid despite the fact that men outperformed women.\textsuperscript{72}

While the manifest relationship standard is quite flexible, it would not permit every physical test to survive scrutiny. Plaintiffs could pre-
vail by demonstrating a viable alternative for measuring the skills sought (strength and endurance) that would not have a disparate im-

pact, but there, of course, the burden would be on the plaintiff.\textsuperscript{73} Also, a physical test designed through an "intuitive process" with no valida-
tion whatsoever will not be sustained, and validation after the fact will not suffice.\textsuperscript{74}

While the selected cutoff scores supposedly designed to measure strength and stamina cannot be arbitrary, any test that measures those skills could be upheld if properly validated.\textsuperscript{75} The emphasis of the

\textsuperscript{65} \textit{Griggs}, 401 US at 432.

\textsuperscript{66} See \textit{Eison v City of Knoxville}, 570 F Supp 11, 13 (E D Tenn 1983) (upholding physical test for police cadets). See also \textit{Robinson v Lorillard Corp}, 444 F2d 791, 798 (4th Cir 1971) (ap-
plying manifest relationship test in context of disparate racial impact).

\textsuperscript{67} 570 F Supp 11 (E D Tenn 1983).

\textsuperscript{68} Id at 12.

\textsuperscript{69} Id at 13. See also \textit{Albemarle}, 422 US at 425.

\textsuperscript{70} \textit{Eison}, 570 F Supp at 13.

\textsuperscript{71} Id ("All proof showed that the exercises were related to physical traits deemed neces-
sary in police officers.").

\textsuperscript{72} The court separately noted that under EEOC Guidelines, the plaintiff could only make out a prima facie case of disparate impact if her police academy class alone was considered the relevant group of applicants. Id.

\textsuperscript{73} Id.

\textsuperscript{74} \textit{Harless}, 619 F2d at 616 (discussing EEOC validation). See also note 18.

\textsuperscript{75} Id. The test in \textit{Harless} required applicants to successfully complete fifteen push-ups, twenty-five sit-ups, a six foot standing broad jump and an obstacle course. Id at 614.
manifest relationship standard is thus placed primarily on the employer's need for the skill in question and a credible demonstration of some relationship between the test and the skill required for successful job performance. While such a standard might be supported by some broad language in Supreme Court precedent, it seems inconsistent with the strict standards often enunciated in the 1970s by the Court and the attempts by the Civil Rights Act of 1991 to reinstate a more pro-plaintiff interpretation of Title VII.

B. The Spurlock Doctrine

A variant of this first approach follows the manifest relationship standard in some contexts, but uses a more pro-employee standard in others. The dividing line is predicated upon whether an employment position implicates safety concerns. While conceding that Congress and the Supreme Court have supported closely scrutinizing most employment tests or requirements, the Spurlock doctrine argues that these bodies have shown more deference to employers where safety is at issue.

Thus, some courts have rationalized that employers should face a lighter burden in proving a "business necessity" when public safety is a concern related to the job position. Some commentators believe a dual standard makes better sense of such Supreme Court precedents as Beazer and Washington that upheld screening mechanisms from disparate impact challenges in safety-sensitive positions. Where the employment test seems calculated to take into account those safety concerns, as most physical tests would, the test will be upheld. These

76 See Griggs, 401 US at 432 ("Congress has placed on the employer the burden of showing that any given requirement must have a manifest relationship to the employment in question."); Washington, 426 US at 250 (upholding literacy test that measured skills "useful, if not essential to satisfactory progress" on the job); Beazer, 440 US at 587 n 31 (finding "manifest relationship" between ban on methadone use and the need for "safety sensitive" employees).

77 See Fitzpatrick v City of Atlanta, 2 F3d 1112, 1119–20 (11th Cir 1993) (expressing concerns for public safety in upholding requirement that all firefighters wear breathing apparatus and remain clean shaven); EEOC v Atlas Paper Box Co, 868 F2d 1487, 1497 (6th Cir 1988) (asserting that courts should be more lenient when evaluating employer disparate impact defenses in jobs affecting public safety); Davis v City of Dallas, 777 F2d 205, 211 (5th Cir 1985) (upholding minimum educational standards as "crucial" to public safety issues when analyzing police employment requirements); Harris v Pan American World Airways, Inc, 649 F2d 670, 675 (6th Cir 1981) (recognizing importance of public safety in airline industry hiring standards); Chrisner v Complete Auto Transit, Inc, 645 F2d 1251, 1262–63 (6th Cir 1981) (requiring lighter burden for "business necessity" in screening of long-distance truckers).

78 See Spiropoulos, 74 NC L Rev at 1479 (cited in note 6) (arguing for a two-tiered treatment of the employer disparate impact defense based in part on public safety concerns).

79 A test might still be struck down if it is based on subjective criteria that do not seem to present a manifest relationship to the safety concern presented. See Banks v City of Albany, 953 F Supp 28, 36 (N D NY 1997) (denying summary judgment where fire department test was based on subjective criteria). Objective tests, like physical ability tests, would usually survive challenge,
public safety concerns will generally involve positions that have a
great deal of contact with and impact upon the public. For example,
the first case to create this bifurcated standard, Spurlock v United Air-
lines, Inc," involved a requirement that applicants for a pilot position
with a major commercial airline log at least five hundred flight hours.8
While this requirement had a profoundly disparate impact on minori-
ties,7 and the defendant produced no evidence that pilots with fewer
than five hundred hours committed more in-flight errors, the court
still held that United Airlines had demonstrated business necessity
and that the requirement was job related.8 The Tenth Circuit ex-
plained that where incalculable “economic and human risks [would
be] involved in hiring an unqualified applicant . . . , the employer bears
a correspondingly lighter burden” than in cases involving unskilled
applicants and no public safety concerns.8 Many of the cases that im-
ply a public safety exception to the strict business necessity defense
have involved either public transportation, as in Spurlock and Beazer,
or police departments, as in Washington. Fire departments, where
many physical tests are also used, would seem to implicate similar is-
sus of public safety and economic and human risk.8 For these rea-
sons, courts evaluating business necessity defenses in the physical test
context have occasionally invoked Spurlock and lightened the burden
on the employer to justify the need for its test.8

While such a standard might make sense of Supreme Court
precedent, the vast majority of these cases were decided prior to the
Civil Rights Act of 1991. A two-tiered standard might be “consistent
with” Supreme Court cases prior to Wards Cove, but there is nothing
in the text of the Act to suggest that the phrase “business necessity”
should involve an entirely different standard in a public safety context

80 475 F2d 216 (10th Cir 1972).
81 Id at 219.
82 Only nine of the airline’s 5,900 flight officers were African-American. Id at 218.
83 Id at 220. Applicants with more than five hundred flight hours were significantly more
likely “to succeed in the rigorous training program which United flight officers go through,” the
components of which were not specified. Id at 218. Still nearly 90 percent of applicants with
fewer hours did complete the program, and there was no evidence in the record of their on-the-
job performance. Id at 219 & n 1.
84 Id at 219.
85 See Zamlen v City of Cleveland, 686 F Supp 631, 654 (N D Ohio 1988) (“The risks to
the public in terms of life, limb and property as a consequence of hiring an unqualified firefighter are
great, resulting in a less heavy burden for demonstrating the job relatedness of the employment
criteria.”), affd, 906 F2d 209 (6th Cir 1990).
86 For jobs requiring minimal skill or human risk, the Spurlock court would require the
employer to bear “a heavy burden to demonstrate . . . that his employment criteria are job-
related.” 475 F2d at 219. While the exact contours are not specified, at least one commentator
has attempted to flesh out a “strict necessity” standard for non-complex jobs. See Spiropoulos, 74
NC L Rev at 1540–46 (cited in note 6).
than it does in other situations, and the doctrine has never been en-
dorsed explicitly by the Supreme Court. On the other hand, dicta in
the Court’s decisions upholding a literacy test for police officers in
Washington and a rule against methadone use among transit employ-
ees in Beazer does lend some support to the idea that the Court im-
plcitly accepted a two-tiered standard. For these reasons, the Spur-
lock doctrine remains controversial.

C. Close Approximation to Job Tasks

A third standard used by several courts in analyzing physical
ability tests focuses on the nature of the test selected by the defendant
employer. When a physical test closely approximates a task the can-
didate would actually perform on the job, it is likely to be upheld. For
example, the court in Hardy v Stumpf was faced with a disparate im-
 pact challenge to an Oakland police department requirement that
each applicant scale a six-foot wall. Because the test “reflect[ed] a
critical duty of Oakland police officers,” it closely approximated a job
task and was determined to be valid.

However, if the test were more of a general fitness exam or an
obstacle course, the court would closely scrutinize the test’s disparate
impact and impose a much higher burden on a department imple-
menting the test. Thus, while the manifest relationship standard and
the Spurlock doctrine focus on the skill being tested, the close ap-
proximation standard looks primarily at the form of the test itself. The
argument in support of this approach appears to be that the best way
to determine if a test is consistent with business necessity is if the em-
ployee would actually be performing the tasks set out for her in the
physical test in the course of her job. Using this standard, courts have

---

87 At least one court has concluded that the Spurlock doctrine is more onerous on the de-
fending employer than Wards Cove. See Zamlen v City of Cleveland, 906 F2d 209, 217 (6th Cir
1990). Thus, the overruling of Wards Cove does not automatically invalidate the Spurlock doc-
tine.

88 See Washington, 426 US at 252 (finding verbal skills to be important in police officer
training); Beazer, 440 US at 587 n 31 (finding that ban on methadone users was job related on
safety grounds).

89 See Smith v City of Des Moines, 99 F3d 1466 (8th Cir 1996) (finding that breathing test
while wearing firefighting apparatus closely simulated job conditions and, therefore, survived at-
tack under the Age Discrimination in Employment Act); Hardy v Stumpf, 21 Cal 3d 1, 576 P2d
1342 (1978) (upholding requirement that police officer applicants successfully scale a six-foot
wall as job related because it significantly correlated to an important element of the job).

90 21 Cal 3d 1, 576 P2d 1342 (1978).
91 Id at 1344.
92 Id at 1346.
93 See Berkman v City of New York, 705 F2d 584 (2d Cir 1983) (affirming district court
finding that battery of physical tests for firefighter positions was not job related); Legault v
aRusso, 842 F Supp 1479 (D NH 1994) (granting preliminary injunction in favor of female fire-
fighter and invalidating four-part agility test including an obstacle course).
upheld a requirement that a police officer scale a six foot wall, or run 1.5 miles in 12 minutes because these are “critical duties” performed in the course of a job. And the New York City Fire Department could implement a test including a hose pull and wall vault that simulated engine and ladder company tasks. On the other hand, a general skills test including a pike pole pull, sandbag drag, and stair climb, was invalid because it did “not represent the physical demands of the job.” Under the close approximation standard, a general battery of physical tests like that approved in *Eison* would presumably be found invalid since it does not measure any actual job tasks.

Courts that require a physical test to closely approximate actual job tasks tend to focus on language in *Albemarle* and *Dothard* that opposed tests failing to meet this standard. The Court in *Dothard* found the height and weight requirements suspect because of a lack of correlation between the skill desired by the Alabama prisons department (strength) and what was actually being measured (weight). If a test closely replicates an actual job task, such as a hose pull or a wall climb, and the employer can demonstrate a rational need for having the employee perform that task, courts are less suspicious that the test is masking invidious discrimination or inadvertently screening out qualified applicants.

The close approximation standard is often tied to the concept of “content validity.” The Equal Employment Opportunity Commission (“EEOC”) has developed guidelines for validating employment tests. One of these measures, content validity, is established by “data showing that the content of a selection procedure is representative of important aspects of performance on the job.” Although the guidelines do not carry the force of legislative enactments, courts have granted them deference. This is especially true in light of the Supreme

---

94 See Hardy, 576 P2d at 1347 (“[I]t is difficult to imagine a more accurate way of testing ability to scale a six-foot wall than to scale one.”); *Lanning v Southeastern Pennsylvania Transportation Authority*, 1998 US Dist LEXIS 9388, *187–88 (E D Pa) (finding that officers need aerobic activity to arrest criminals and that requiring potential officers to run 1.5 miles in twelve minutes appropriately measured this need), rev’d, 181 F3d 478 (3d Cir 1999).

95 See *Berkman v City of New York*, 812 F2d 52, 54 (2d Cir 1987).

96 *Brummitt v City of Columbus*, 642 F Supp 1214, 1250 (S D Ohio 1986). See also *Russo*, 842 F Supp at 1489 (finding that the fire department failed to validate an obstacle course as being job related).

97 See *Eison*, 570 F Supp at 12 (describing approved fitness test consisting of push-ups, pull-ups, sit-ups, and two-mile run).

98 See 433 US at 331 (“[T]he appellants produced no evidence correlating the height and weight requirements with the requisite amount of strength thought essential to good job performance.”).


100 See, for example, *Association of Mexican-American Educators v California*, 195 F3d 465, 486 (9th Cir 1999) (stating that noncompliance with EEOC guidelines diminishes the probative value of a defendant’s validation study); *Allen v Entergy Corp*, 181 F3d 902, 905 (8th Cir 1999)
Court’s decision to generally require statistical proof of a test’s validity.\textsuperscript{1}

For these reasons, courts have been more sympathetic to a test replicating job conditions than a mere “fitness test” because the former measures the component skills that are important to the job. While courts using this standard generally inquire solely into the relation between the test and the skill being measured, they may still reject the selection of an arbitrary cutoff score that has no rational relation to the needs of the job,\textsuperscript{102} or that has been positively demonstrated to be unnecessarily high.\textsuperscript{103}

D. Minimum Qualifications

The strictest standard to date was recently announced by the Third Circuit in \textit{Lanning v Southeastern Pennsylvania Transportation Authority}.\textsuperscript{104} At issue was a requirement that all applicants for police officer positions assigned to the transit authority for the greater Philadelphia metropolitan area (“SEPTA”) successfully run 1.5 miles in under twelve minutes.\textsuperscript{105} The test had been designed by a fitness consultant in response to high crime rates within the transit system and several incidents in which out-of-shape officers were injured or arrived late to crime scenes.\textsuperscript{106} Police officers for the transit authority usually worked alone, patrolled on foot, and were expected at times to pursue suspects.\textsuperscript{107} The test was designed to measure a specific aerobic capacity that the fitness consultant and transit authority determined would be necessary for police officers to have a good chance of apprehending the average criminal.\textsuperscript{108}

When the test was implemented, a total of 55.6 percent of men were able to pass the test, but only 6.7 percent of women ran the 1.5

\textsuperscript{101} See \textit{Albemarle}, 422 US at 431, quoting 29 CFR § 1607.4(C) (requiring employer to show a written skills test was “significantly correlated with important elements of work behavior”).

\textsuperscript{102} See \textit{Pietras v Board of Fire Commissioners}, 180 F3d 468, 472 (2d Cir 1999) (invalidating requirement that all firefighter candidates drag a hose 150 feet in under four minutes after finding no rational basis for time selected).

\textsuperscript{103} See \textit{EEOC v Simpson Timber Co}, 1992 US Dist LEXIS 5829, *10–11 (W D Wash) (rejecting cutoff score used in physical test by timber company after finding that a lower cutoff level would reduce disparate impact while still measuring job qualifications).

\textsuperscript{104} 181 F3d 478 (3d Cir 1999).

\textsuperscript{105} Id at 482.


\textsuperscript{108} \textit{Lanning}, 181 F3d at 482.
miles successfully.\textsuperscript{109} The transit authority conceded that the distinction in percentage passage rates was enough to create a prima facie case of disparate impact.\textsuperscript{110} Under the Civil Rights Act of 1991, the burden of proving that the test was a business necessity then shifted to the defendants.\textsuperscript{111}

The district court, relying on the close approximation to job tasks test, held that the defendant had proved that the test was a business necessity.\textsuperscript{112} It accepted the testimony of the fitness consultant regarding the need for physically fit transit police officers, and the ability of the test to measure that skill. Accordingly, the court granted summary judgment for the defendant.\textsuperscript{113}

The Third Circuit reversed the district court's decision and remanded the case for further discovery.\textsuperscript{114} The court emphasized the most stringent language appearing in the Supreme Court's Title VII jurisprudence.\textsuperscript{115} The primary justification for this emphasis was the fact that the Civil Rights Act of 1991 was a direct response to \textit{Wards Cove}, which had articulated an easier standard for employers to meet.\textsuperscript{116} Consequently, the court found that Congress wished to eliminate employment requirements with a disparate impact except in very narrow circumstances and determined that a strict standard was needed to accomplish this aim: "[A] discriminatory cutoff score is impermissible unless shown to measure the minimum qualifications necessary for successful performance of the job in question."\textsuperscript{117} While acknowledging the strictness of this standard, the court did not believe it would require the implementation of quota systems or the hiring of unqualified applicants. Instead, the court articulated three non-quota

\begin{itemize}
\item \textsuperscript{109} Id at 483.
\item \textsuperscript{110} Id.
\item \textsuperscript{111} See 42 USC § 2000e-2(k)(1)(A) (codifying the employer disparate impact defense).
\item \textsuperscript{112} \textit{Lanning}, 1998 US Dist LEXIS 9388 at *147-48 (holding that employer disparate impact defense allows "employment practices that significantly serve, but are neither required by nor necessary to, the employer's legitimate business interests").
\item \textsuperscript{113} Id at *208-09.
\item \textsuperscript{114} \textit{Lanning}, 181 F3d at 494.
\item \textsuperscript{115} Id at 489. See also \textit{Albemarle}, 422 US at 434 ("The fact that the best of those employees . . . score well on a test does not necessarily mean that that test, or some particular cutoff score on the test, is a permissible measure of the minimal qualifications of new workers entering lower level jobs."); \textit{Dothard}, 433 US at 331-32 n 14 ("[A] discriminatory employment practice must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge.").
\item \textsuperscript{116} \textit{Lanning}, 181 F3d at 488 ("By Congress' distinguishing between \textit{Griggs} and \textit{Wards Cove}, we must conclude that Congress viewed \textit{Wards Cove} as a significant departure from \textit{Griggs}. Accordingly, because the Act clearly chooses \textit{Griggs} over \textit{Wards Cove}, the Court's interpretation of the business necessity standard in \textit{Wards Cove} does not survive the Act.").
\item \textsuperscript{117} Id at 489. Compare 29 CFR § 1607.5(H) (1999) (requiring only that cutoff scores "be reasonable and consistent with normal expectations of acceptable proficiency within the work force") (emphasis added).
\end{itemize}
options available to SEPTA. First, SEPTA could use scores on the test not as a hiring device but as an incentive program. Second, it could "validate a cutoff score ... that measures the minimum [qualifications] necessary to successfully perform the job." Third, SEPTA could utilize two different cutoff scores—one for men and an easier one for women. Because the district court utilized a different standard, the case had to be remanded for evaluation of the 1.5 mile run under the test laid down by the Third Circuit.

Dissenting, Judge Weis argued that a close examination of the Supreme Court's Title VII jurisprudence before Wards Cove showed great deference to employer decisions when public safety was implicated. He feared that the majority's new minimum qualifications test would lead to the abandonment of all physical tests and concluded that the overriding consideration in these cases was the issue of public safety endorsed in the Spurlock doctrine. Since public safety concerns are implicated in police department hiring decisions, Judge Weis would have granted summary judgment to SEPTA if it could demonstrate a manifest relationship between its test and the job requirements, a standard more pro-employer even than that used by the district court.

The debate between the majority and the dissent in Lanning is only another example of the difficulties courts face when evaluating the use of physical fitness tests as a cutoff for employment purposes. Because the Supreme Court has yet to interpret this portion of the 1991 Civil Rights Act, courts must struggle to resolve these confused standards on their own.

118 Lanning, 181 F3d at 490 n 15.
119 Id.
120 Id. The court's reference to validation seems to refer to the validation procedures found in the EEOC guidelines. However, it is unclear how an employer could validate a minimum cutoff score under current guidelines, which focus on the qualifications of the test and testing procedures, and not on any actual score used as the cutoff.
121 Id. This third option might be prohibited by the Civil Rights Act of 1991, which prevents the "use [of] different cutoff scores for ... employment related tests on the basis of ... sex." 42 USC § 2000e-2(i) (1994).
122 Lanning, 181 F3d at 494.
123 See, for example, Albemarle, 422 US at 434 (allowing employment tests that "fulfill a 'genuine business need'"); Dothard, 433 US at 320 (concluding that an "employer must show that a requirement has 'a manifest relationship to the employment in question'"), quoting Griggs, 401 US at 432; Beazer, 440 US at 587 n 31 (allowing a test that "significantly served" the goals of safety and efficiency, even if the test was not required for those goals to be served).
124 See Lanning, 181 F3d at 502 (Weis dissenting).
125 Id at 500.
IV. A PROPOSED RESOLUTION

In evaluating the business necessity defense after 1991, it is crucial to scrutinize the text of the statute Congress enacted. While canons of statutory construction counsel interpretation of statutes in light of the prevailing common law, any such interpretation must still make sense of the current statutory text. That text contains two requirements: the test must be "job related" and it must be "consistent with business necessity." While most courts essentially adopt a single standard to conflate both of these terms, it is possible, and indeed makes sense, to treat them as two separate requirements, both of which must be met for an employer to successfully defend the validity of its test.

The best way to make sense of the statutory text and case law would be to create a two-prong test for assessing physical ability tests and other employment screens that create a disparate impact, one addressing the "job related" requirement and the other assessing the "business necessity" component. It should be possible for an employment test to be struck down under either component without automatically implicating the other. At the same time, it should be possible for employers to muster proof that a given test satisfies both prongs and thereby persuade a court that the test should be sustained in spite of its disparate impact.

A. The Two-Prong Test

1. The skill sought to be measured by the employment test is consistent with business necessity.

A two-prong test also makes sense because most cases evaluating employment tests emphasize one of two separate concerns. The first is whether the skill being measured (for example, speed, stamina, strength, intelligence) by the employment test is actually a prerequisite for the job in question. This concern reflects whether or not the test is "consistent with business necessity" by measuring a skill necessary for job performance. In order to meet the demands of this first prong, an employer would need to identify the skill or ability the employment test measures, and prove to the district court's satisfaction a "manifest relationship" between the skill tested and the job itself, evidence similar to that amassed in Eison.

\[\text{\cite{Eison}}\]

Canons of construction may be even more crucial in interpreting this section of the Civil Rights Act of 1991 because usage of the other major source of statutory interpretation, legislative history, is largely foreclosed by the text of the Act itself. See Civil Rights Act of 1991, Pub L No 102-166 § 105(b), 105 Stat 1071, 1075 (1991).

\[\text{\cite{2000e-2(k)(1)(A)(i)}}\]
In the context of physical ability tests, most employers would be able to pass this prong since it is similar to the manifest relationship standard outlined in *Eison* and related cases and could be satisfied through the production of similar evidence of a job’s physical demands. Physical skills such as strength, speed, and endurance are generally rationally related to police and fire department positions. Of course, physical tests might not be manifestly related to a private sector job, or a position as a police dispatcher. And some tests would fail to satisfy even the manifest relationship standard.

The first prong makes sense because often the first issue that has confronted courts in employment discrimination cases is whether the skill tested is actually necessary for the job. Thus, the defendant in *Griggs* offered no evidence whatsoever that literacy or basic math skills, which might possibly be measured by the receipt of a high school diploma, were necessary for the position of a coal handler at a power plant. Literacy was also not shown in *Albemarle* to be a prerequisite for employment as a low-level worker at a paper mill. On the other hand, the methadone rule in *Beazer* did significantly serve the goal of a safe and accident-free transit authority. The Court’s primary concern in these cases seemed to be that employers were actually using these neutral criteria not to screen out unqualified applicants but to conceal disparate treatment of protected minorities. Meeting the first prong by proving the business necessity of the skill in question would serve to rebut such fears.

2. The test itself is clearly job related by closely approximating an on-the-job task.

The first prong is a necessary, but not a sufficient, condition for a successful business necessity defense. Equally important is the requirement that the test selected by the employer accurately gauge that skill. Thus, if a given skill is found to be a “business necessity” for a job, then a test that precisely measures that skill is “related” to the job. Consider the decision in *Dothard*. The Court left open the possibility that possession of some amount of strength might be a bona fide prerequisite to employment as a prison guard, which would render that

128 See *Zamlen v City of Cleveland*, 906 F2d 209, 218 (6th Cir 1990) (accepting expert testimony that firefighters must possess high degrees of “aerobic capacity and aerobic fitness, . . . muscle strength, muscular endurance, flexibility, coordination, muscle balance, and speed”); *Berkman v City of New York*, 812 F2d 52, 59–60 (2d Cir 1987) (finding that “firefighters frequently face situations where their anaerobic abilities determine whether or not they will save the lives of the fire victims”).

129 See notes 73–74 and accompanying text.

130 See *Griggs*, 401 US at 431–32.

131 See *Albemarle*, 422 US at 431–32.
skill a “business necessity” under the first prong of the proposed test. However, the Court required that any test must correlate to “the requisite amount of strength thought essential to good job performance.”

A simple height requirement at best indirectly correlates with strength, as short individuals may in many instances be stronger than those who are taller. A direct measurement is likely to be most successful if it measures the actual content of the job itself. This “content validation” is one of the three mechanisms that the EEOC suggests can be used to validate an employment test, and it is best used for skills that are easily measurable (as opposed to abstract thinking skills or leadership). Since physical strength and stamina are concrete skills, it should not be unduly burdensome for an employer to devise tests that closely approximate job tasks. Thus, in the physical ability test arena, “job relatedness” is likely to become synonymous with the language “closely approximates a job task” found in numerous opinions over the years. In addition, an employer would probably have to show that the cutoff score is also rationally calibrated to approximate necessary on-the-job skills, although that score need not be the absolute minimum required for job performance.

To summarize, the best interpretation of the business necessity defense to a disparate impact claim codified in the Civil Rights Act of 1991 would uphold the test only when the employer produces evidence that its test closely approximates and effectively measures (“job related”) a skill that is important to successful job performance (“consistent with business necessity”).

132 See Dothard, 433 US at 331–32 (accepting possibility that state could validate a test that measured strength directly, even if such a test created a disparate impact).

133 Id at 331.

134 See 29 CFR § 1607.16(E)–(F) (1999). See also 20 CFR § 1607.14(C)(1) (1999) (stating that content validity studies are inappropriate for measuring such skills as “intelligence, aptitude, personality, commonsense, judgment, leadership, and spatial ability”).

135 Devising such a test would not be costless of course. Since the Civil Rights Act places the burden of proof on the defendant, any employer would be well advised to secure an expert to devise the test and validate it in accordance with EEOC procedures. A mere assertion of job relatedness “bereft of any evidence” is still unlikely to survive judicial scrutiny. See Pietras v Board of Fire Commissioners, 180 F3d 468, 472 & n 5 (2d Cir 1999) (invalidating timed hose pull in the face of total lack of evidence that time chosen “reflected the needs of the job”).

136 See Fitzpatrick v City of Atlanta, 2 F3d 1112, 1120 (11th Cir 1993) (upholding test measuring breathing capacity while wearing firefighter's mask); Zamlen v City of Cleveland, 906 F2d 209, 218–19 (6th Cir 1990) (accepting a hiring test that simulated actual firefighting tasks); Hardy, 576 P2d at 1342 (upholding the requirement that police cadets climb a six-foot wall as reflecting important police skills).

137 See Pietras, 180 F3d at 471 (invalidating cutoff score that was the average of all previous test-takers and not related in any way to job performance).
B. Benefits of the Two-Prong Test

1. Consistent with statutory text.

An important canon of construction requires courts to give meaning to the entire statutory text if possible. If there are two possible readings of a statute, one of which "renders some words altogether redundant," the other reading should be preferred. It is presumed that Congress knows what it is doing in enacting statutory text and does not include words that are unnecessary for understanding the statute.

The best interpretation of the business necessity defense under the standard would be one that gives meaning both to the phrase "job related" and to the phrase "consistent with business necessity." Indeed, at least one court favoring a strict standard has argued that the manifest relationship standard violates this principle by requiring only that the physical ability test be related to the job in question without evaluating the business necessity of the practice.

At the same time, it could equally be argued that a stringent minimum qualifications test creates the same problem by writing out the job relatedness standard at the expense of the business necessity language. If "consistent with business necessity" defines a test that measures the minimum qualifications necessary for successful job performance, it seems impossible to conceive of a test that would meet such a standard and yet not be job related. In that case, the "job relatedness" prong of the business necessity defense would become irrelevant.

The two-prong test is more consistent with statutory text than the options currently outlined by courts because it recognizes the inherent difficulty in applying a single standard to cover the two separate terms "job related" and "consistent with business necessity." By bifurcating these terms, focusing the attention of one on the job relatedness of the employment test itself and the other on the business necessity of the

---

138 See United States v Gonzalez, 520 US 1, 8 (1997) (giving meaning to entire statutory text by disregarding legislative history); Walters v Metropolitan Educational Enterprises, Inc, 519 US 202, 209 (1997) (requiring that each word of a statute be given meaning); United States v Nordic Village, Inc, 503 US 30, 36 (1992) ("[A] statute must, if possible, be constructed in such fashion that every word has some operative effect.").


140 McNary v Haitian Refugee Center, Inc, 498 US 479, 496 (1991) (presuming that Congress has knowledge of basic rules of statutory construction); John F. Manning, Constitutional Structure and Judicial Defeference to Agency Interpretations of Agency Rules, 96 Colum L Rev 612, 625 (1996) ("[T]he Court routinely presumes that Congress enacts statutes against an interpretive background that includes the Court's norms of construction.").

141 See Lanning, 181 F3d at 489 ("Judicial application of a standard focusing solely on whether the qualities measured by an entry level exam bear some relationship to the job in question would impermissibly write out the business necessity prong of the Act's chosen standard.").
skill being tested, each can be given statutory meaning without caus-
ing "job related" to be swallowed up or "consistent with business ne-
ces sity" to be ignored.

In addition, it is crucial to note that the Civil Rights Act requires
only that an employment practice be consistent with business neces-
sity. Implementing a standard like the minimum qualifications test
requires the test to, in essence, be a business necessity, writing the
word "consistent" out of the statute. Thus, in evaluating whether a test
is consistent with business necessity, the focus must lie not on the test
itself, as it is under the minimum qualifications standard, but on the
skills being measured by the test. If those skills are consistent with the
needs of a safe and efficient business, the test is valid.

At the same time, the proposed two-prong test clearly goes far-
ther than either the manifest relationship standard or the Spurlock
document. The Spurlock exception, while containing many valid con-
cerns about public safety, is not supported by the text of the statute;
there is simply no distinction made on the statute's face between pub-
lic and private employers or between jobs that do and do not impact
public safety. While many of these safety concerns can still be evalu-
ated when assessing the skills necessary for job performance, they
should not be taken into account as independent considerations.

2. Consistent with Supreme Court precedent.

The interpretive memorandum accompanying the Civil Rights
Act of 1991 explicitly requires that the Act be applied consistent with
Supreme Court decisions predating Wards Cove. Therefore, no in-
terpretation of Title VII's current statutory language is acceptable if it
does not reflect the meaning of the terms "business necessity" and
"job related" as they were articulated in those precedents. Any test
must, at a minimum, both correctly incorporate the five major Su-

\[143\] See Beazer, 440 US at 587 & n 31 (upholding rule banning methadone users from em-
ployment in "safety sensitive" positions). Just as the skill or qualification in Beazer was safe job
performance, so in physical tests the skill of strength or stamina should be judged based on the
needs of the police or fire department in question.
\[144\] See text accompanying note 87.
\[145\] See Part IV.B.4.
\[146\] By endorsing the manifest relationship standard in the public safety context, Spurlock
also retains that standard's shortcoming of essentially writing the "business necessity" compo-
nent out of the statute. See text accompanying note 141.
\[147\] See Interpretive Memorandum, 137 Cong Rec S 28680 (Oct 25, 1991).
\[148\] While dicta in Wards Cove itself found the Court's precedents to endorse a manifest re-
lationship or Spurlock standard, see Wards Cove, 490 US at 659–60, that analysis cannot bear on
interpretation of the Act, which explicitly orders interpreting courts to rely on decisions prior to
Wards Cove. See 137 Cong Rec S 28680.
supreme Court Title VII precedents and harmonize their reasoning. The two-prong test does this by recognizing that Griggs and its progeny were actually focused on two different problems in their language and rhetoric: the problem of employers testing skills that are unnecessary for job performance (i.e., business necessity) and the problem of tests that incorrectly measure necessary skills (i.e., job relatedness). Thus, when the Court notes that “[t]he touchstone is business necessity” and that a practice must be necessary for safe and efficient performance, it is mainly focusing on the skills being tested: literacy among coal handlers or strength among prison guards. But when the court seeks a “demonstrably . . . reasonable measure of job performance,” it is questioning whether the test itself is a close approximation or an adequate measure for that skill.

3. Consistent with purposes of Title VII.

The primary goal of Title VII was to prevent discriminatory hiring practices. While there is much debate over the exact scope and purpose of disparate impact analysis, at a minimum, disparate impact analysis should halt practices that appear likely to mask discriminatory treatment or that do not produce any countervailing benefits to the employer. An employment test that fails to measure accurately a job’s requirements and that creates a disparate impact provides an employer with no legitimate benefits or efficiencies and harms society by reducing job opportunities for females or minorities. If the courts can successfully identify such cases, they should act to rectify the problem.

In the context of physical ability tests, the two-prong test does a good job of addressing the purposes of Title VII on two fronts. First, it is unlikely that an employer could select a physical test with discriminatory intent that would both achieve its aims and survive judicial scrutiny. An employer could not simply choose to test skills at which

150 Griggs, 401 US at 431.
152 Dothard left open the possibility that strength might be a valid business necessity for employment as a prison guard but found that the test selected to measure strength (height and weight requirements) was not “job related.” See 433 US at 331–32.
154 See Spiropoulos, 74 NC L Rev at 1485 (cited in note 6) (noting that the objectives of Title VII were to remove “artificial barriers” to employment).
156 See Spiropoulos, 74 NC L Rev at 1528–29 (cited in note 6) (arguing that Title VII was intended to balance employer needs while preventing discrimination).
men excel if those skills are not a business necessity. Alternatively, the employer could not create a test slanted to benefit males. If the employer created a test that did not accurately measure a skill necessary for job performance, that test would be struck down as not job related. And if an attempt were made to manipulate cutoff scores to avoid hiring female employees, the test again would be invalid under the job relatedness prong because the placement of the cutoff score would not be rationally related to the skill being tested.

Even if disparate impact analysis is also thought necessary to prevent hidden, subconscious stereotypes from unfairly discriminating against women, the two-prong test will still likely remedy the problem. Since the employer must in any event demonstrate content validity as to an employment test's make-up, any test implemented without forethought will have virtually no chance of viability.

Second, this test is likely to reduce error costs since it both clarifies and limits the inquiry courts must make. A disparate impact inquiry can create error costs in either of two directions. Courts might underenforce the statute, permitting discriminatory practices and reducing confidence in public agencies. On the other hand, courts might overenforce the statute, leading to costly validation procedures or the selection of unqualified applicants, which, in turn, would reduce the effectiveness of law enforcement or fire protection.

The two-prong test balances these concerns by requiring judges to seek answers to questions they are likely to resolve correctly. It may be difficult for a court to determine, based on competing statistical displays, whether a police officer must run a mile in eight minutes, ten minutes, or fifteen minutes in order to successfully perform her job, as would be required under the minimum qualifications test. It should be much less difficult for a court to determine (a) that endurance is (or is not) a critical aspect of employment as a police officer and (b) that a 1.5 mile test does (or does not) approximate job conditions the aver-

---

157 See Berkman v City of New York, 705 F2d 584, 588 (2d Cir 1983) (upholding district court finding that physical test that emphasized anaerobic and male-prefering skills was invalid under Title VII).
158 See Legault v aRusso, 842 F Supp 1479, 1488 (D NH 1994) (finding that four-part agility test did not test job skills based on their relative importance to the job of firefighter).
159 See Pietras v Board of Fire Commissioners, 180 F3d 468, 472 (2d Cir 1999) (blocking arbitrary selection of hose drag time). In addition, after-the-fact manipulation would probably create evidence of disparate treatment.
161 See Colker, 19 UC Davis L Rev at 772–75 (cited in note 60) (discussing historical gender-based segregation in police and fire departments); Flynn, Despite Recruiting, Few Women Do Well, NY Times at A1 (cited in note 61) (discussing efforts to improve hiring of females in New York City fire department); David Holmstron, Women Officers Arrest the Gender Gap, Christian Science Monitor 11 (Jan 12, 2000) (noting that presence of women may improve police forces).
Physical Ability Tests

2000]

The two-prong test is thus not only consistent with the statutory mandate and Supreme Court precedent, but also with easy implementation and prevailing notions of judicial competence.

4. Unlikely to compromise public safety.

While the proposed test does not endorse the Spurlock doctrine's emphasis on public safety and efficiency, it does essentially incorporate safety and efficiency concerns through the "business necessity" prong of the employer defense. For example, a police department would still be allowed to implement a running requirement clearly motivated by concerns about criminals eluding pursuit and embarrassing media attention. In effect, the test does grant some deference to employers in seeking to justify their employment practices if they are testing skills that will create the safest and most efficient police and fire departments.

C. Application of the Proposed Test

1. Within the physical test context.

Under the proposed two-prong test, most of the cases discussed in Part III of this Comment would still be decided in the same way. For example, the Hardy court would merely determine under the first prong that strength and climbing skills were necessary for the job of an Oakland police officer and then recognize, as it did, that a test requiring climbing a six-foot wall is job related because it closely approximates those tasks. In a few cases, however, this refined analysis would yield different results. The court in Eison, for example, would recognize that while physical fitness was a "business necessity" for a police officer, measuring that skill through push-ups, sit-ups, and an obstacle course that do not closely approximate actual job tasks creates a test not sufficiently job related to withstand Title VII scrutiny. A case like Lanning might also be decided differently under the two-prong test. While the court there required that the test itself be a business necessity, the two-prong test would probably deem endurance running to be a skill necessary for successful job performance. Based

162 To sustain such a test, a court could rely either on internal department records that such runs occurred on the job or testimony from experts conducting a job analysis that such a run would simulate actual job conditions. See 29 CFR § 1607.14(C)(2) (1999) (describing technical components and standards of a job analysis).

163 Indeed, Lanning appears to be one of the few physical test cases where clear evidence indicated the test was not established for any discriminatory purpose. Compare Legault v Zambarano, 105 F3d 24, 26–28 (1st Cir 1997) (affirming Rule 11 and discovery sanctions against attorneys and defendants for concealing that invalidated physical test was a "mirage").
on the validation tests entered into evidence by SEPTA, a court would likely conclude that a 1.5 mile run closely approximates such actual job tasks as officer backups and running assists and that the twelve minute cutoff time was not selected arbitrarily. The proposed test improves upon the minimum qualifications test by recognizing that it is only the skill tested that must be “consistent with business necessity” while the testing mechanism selected must be “job related.”

2. Outside the physical test context.

Finally, while the focus of this Comment has been on establishing a business necessity test and its ramifications in the context of employer physical ability tests, the two-prong test proposed here could easily be applied to all Title VII disparate impact claims. Most other challenged employment screens are either general qualifications requirements or written tests. In either of these cases, the basic analysis would begin by questioning whether the skill the employer claimed to be testing was one that would be necessary for safe and efficient job performance.

At the second stage, the court would then inquire whether the test itself was “job related” in that it accurately measured the skill desired. The inquiry might be slightly different in the physical and non-physical contexts, but, in either case, the EEOC guidelines could provide guidance for test validation. In many cases, the question would be identical to that found in the physical ability context: Does the test closely approximate an actual job task? If so, the test will survive judicial scrutiny. Even if it did not, some higher-order thinking tests might still be upheld through construct or criterion validation. The principles behind the two-prong test, however, would remain the same.

164 Each of the two hundred SEPTA officers averaged 9.6 backups (run as paced jogs) and 1.9 running assists (run as fast as possible while maintaining sufficient reserve energy) per year. See Lanning v Southern Pennsylvania Transportation Authority, 1998 US Dist LEXIS 9388, *13–14 (E D Pa), revd, 181 F3d 478 (3d Cir 1999).


166 See id. See also Williams v Ford Motor Co, 187 F3d 533, 539–41 (6th Cir 1999) (noting that criterion studies examine “whether performance on [a] test adequately correlates with performance on the job” in holding that a battery of mental and physical tests was both content and criterion valid). Construct validity studies are best used for difficult-to-test but important traits such as creativity and require extensive validation. See Zamlen v City of Cleveland, 906 F2d 209, 218 (6th Cir 1990); 29 CFR § 1607.14(D).
CONCLUSION

The courts and Congress alike have struggled with Supreme Court precedent formulating the business necessity defense because of the frequent failure to differentiate between the two different strands of analysis that have gone into that defense: business necessity and job relatedness. Courts should utilize a test that accounts for both of these strands. When an employer defends a physical test that has a disparate impact on a protected class, the employer should be required to demonstrate both that it is attempting to test a skill consistent with its business necessity and that it has selected a test that closely approximates actual job-related tasks.