NOTES

TITLE VII DISCRIMINATION IN BIOCHEMICAL TESTING FOR AIDS AND MARIJUANA

The acquired immune deficiency syndrome (AIDS) and drug abuse in the workplace are the subjects of relentless media and government scrutiny.[1] As of early 1987, public health officials had reported over 30,000 cases of AIDS in the United States.[2] By the end of 1987, at least one million Americans probably had been exposed to the fatal disease.[3] The associated health insurance costs[4] and productivity losses[5] will have

1. *See, e.g.*, PRESIDENT'S COMM'N ON ORGANIZED CRIME, AMERICA'S HABIT: DRUG ABUSE, DRUG TRAFFICKING AND ORGANIZED CRIME (1986); SURGEON GENERAL'S REPORT ON ACQUIRED IMMUNE DEFICIENCY SYNDROME (1986) (providing current medical information to the American public); *The AIDS Threat: Who's at Risk?*, NEWSWEEK, Mar. 14, 1988, at 42 (outlining Drs. Masters, Johnson and Kolodny's pessimistic projections regarding of the spread of AIDS into the heterosexual community); *The AIDS Conflict*, NEWSWEEK, Sept. 23, 1985, at 18 (describing how ignorance about the disease is causing an epidemic of fear in the population); *Officials in Discord Over U.S. Proposal for AIDS Virus Test*, N.Y. Times, Feb. 5, 1987, at A19, col. 1 (Federal proposal to administer test for AIDS virus to anyone admitted to a hospital or applying for a marriage license causes confidentiality concerns.); Altman, *New Fear on Drug Use and AIDS*, N.Y. Times, Apr. 6, 1986, at 1, col. 2 (Intravenous drug use is prime cause of AIDS infection among heterosexuals in America.).

2. *See* Francis & Chin, *The Prevention of Acquired Immunodeficiency Syndrome in the United States: An Objective Strategy for Medicine, Public Health, Business, and the Community*, 257 J. A.M.A. 1357, 1357 (1987).

3. *Id.*

4. Insurance companies contend that people exposed to the AIDS virus "are likely to have higher health care costs" than the rest of society. Brandt, *AIDS: From Social History to Social Policy*, 14 LAW MED. & HEALTH CARE 231, 237 (1986). Claire Wolkoff of the American Academy of Actuaries summarized the insurance industry's position on AIDS testing: "If America's private voluntary-insurance system is to remain workable, AIDS tests must be allowed so the disease can be underwritten in the same manner as heart disease, cancer, or alcohol and drug abuse. . . . The alternative is to spread the risk factor over the whole population, thus raising the price of insurance for everyone." *Remove Stigma From AIDS Test: An Insurer's View*, N.Y. Times, June 11, 1986, at A34, col. 4.

As the number of AIDS cases and the cost of treatment increase, there may be growing pressure on state governments to allow insurers to test for the human immunodeficiency virus (HIV) antibodies and to exclude carriers from coverage. *See* Gostin, *The Nucleus of a Public Health Strategy to Combat AIDS*, 14 LAW MED. & HEALTH CARE 226, 229 (1986); Comment, *AIDS and Insurance: The Rationale For AIDS-Related Testing*, 100 HARV. L. REV. 1806, 1814-17 (1987).

5. The loss of productivity of people exposed to HIV will vary depending on the stage of the infection. Patients with AIDS-Related Complex (ARC), infected with HIV but not suffering from full-blown AIDS, suffer from specific signs and symptoms. These diagnostic markers include "persistent fevers, diarrhea, anorexia (loss of appetite), weight loss, and malaise not due to other underlying illnesses" as well as progressive generalized lymphadenopathy (enlarged lymph glands that have

a significant impact on the private sector. Likewise, the Research Triangle Institute estimates that substance abuse on the job costs the economy as much as forty-six billion dollars in 1980.[6] Federal experts estimate that between ten percent and twenty-three percent of workers use drugs on the job.[7]

In response to mounting publicity and concern, increasing numbers of private employers have implemented biochemical testing to screen employees for drug use and for the presence of AIDS.[8] Employers and insurers will probably also make great use of similar tests for AIDS in coming years. Courts and employers, however, should interpret these biochemical tests with caution. Although facially neutral, biochemical tests may discriminate against groups protected by Title VII of the Civil Rights Act of 1964.[9]

Title VII, in addition to banning intentional discrimination, prohibits facially neutral employment practices that have a disparate impact on

---

grown progressively larger in scattered areas of the body). Mayer, *The Clinical Challenges of AIDS and HIV Infection,* 14 LAW MED. & HEALTH CARE 281, 282 (1986).

Once full-blown AIDS develops, the patient may suffer from a variety of opportunistic infections, which take advantage of the patient's depressed immune system. The most common of these infections is the *Pneumocystic carinii* pneumonia, the most frequent cause of death among AIDS patients. Opportunistic neoplasms such as Kaposi's sarcoma and certain lymphomas also may impair worker health. *Id.* at 282-83.

Another severe impairment of productivity will come from the development of HIV neurologic disease ("AIDS encephalopathy"). HIV proliferates in nerve cells of the brain, spinal cord and peripheral nerves, causing a variety of neurologic diseases. *Id.* at 283-84; *see also* Snider, Simpson, Nielson, Gold, Metroka & Posner, *Neurological Complications of Immune Deficiency Syndrome: Analysis of 50 Patients,* 14 ANNALS OF NEUROLOGY 403, 403-04 (1983). The frequency of this progressive encephalopathy, in which "dementia and motor dysfunction predominate," is "presently uncertain, but it may eventually afflict the majority of AIDS patients." Navia, Cho, Petito & Price, *The AIDS Dementia Complex: II Neuropathology,* 19 ANNALS OF NEUROLOGY 525, 525 (1986); *see* Epstein, Sharer, Joshi, Fojas, Koenigsberger & Oleske, *Progressive Encephalopathy in Children with Acquired Immune Deficiency Syndrome,* 17 ANNALS OF NEUROLOGY 488, 496 (1985) ("progressive encephalopathy may represent an early feature of AIDS in children"). Early findings of the dementias include impaired memory and visual-spacial orientation. Mayer, *supra,* at 283-84.

6. H. HARWOOD, D. NAPOLITANO, P. KRISTIANSEN & J. COLLINS, ECONOMIC COSTS TO SOCIETY OF ALCOHOL AND DRUG ABUSE AND MENTAL ILLNESS: 1980, at 3 (Research Triangle Inst. 1984); *see also* Cohen, *Drugs in the Workplace,* 45 J. CLINICAL PSYCHIATRY, Dec. 1984, pt. 2, at 4, 4 (1984) ("Metropolitian Life Insurance Company estimates the direct costs to industry at $85 billion a year.").

7. *Battling the Enemy Within: Companies Fight to Drive Illegal Drugs Out of the Workplace,* TIME, Mar. 17, 1986, at 53.

8. *See The Ruckus Over Medical Testing,* FORTUNE, Aug. 19, 1985, at 57 (percentage of Fortune 500 companies screening applicants rose from 10% in 1982 to nearly 30% in 1985); Noel Dunivant & Assocs., Drug Testing in Major Corporations: A Survey of the Fortune 500, at 4, 6 (October, 1985) (conducted for Compuchem Laboratories) (finding that of 180 Fortune 500 firms interviewed, 18% acknowledged testing procedures and another 19% stated that they anticipated the implementation of biotesting within the next two years).

9. The Civil Rights Act of 1964, §§ 701-718, 42 U.S.C. §§ 2000e to 2000e-17 (1982).

persons of a particular race, national origin, sex, or religion.[10] Under the disparate impact theory, a plaintiff need only prove a significant disparate impact to establish a prima facie claim; an intention to discriminate is not a necessary element.[11] Once the plaintiff meets this burden, the employer must justify the employment practice.[12]

Judicial decisions dealing with disparate impact may be separated into two categories: those dealing with selection tests and those dealing with selection policies. This distinction is important: if biochemical testing is classified as a selection policy, an employer need only articulate a single nondiscriminatory reason for the policy. Selection tests, however, are given more stringent scrutiny by reviewing courts.[13] Because biochemical tests are to some degree inherently inaccurate, it is possible that they discriminate on a chemical level.[14] Thus, biochemical tests have the same potential to discriminate as other professionally developed tests. Consequently, biochemical tests are most appropriately analyzed under the testing strand of disparate impact doctrine.

This note argues that a disparate impact analysis must distinguish between the accuracy of a biochemical test and the policy the test may be intended to enforce.[15] An employer seeking to justify a discriminatory biochemical test should follow a two-part method of validation. First, the employer should offer specific evidence verifying the accuracy of the test. The best way to ensure the accuracy of biochemical testing is to follow initial screening tests with a confirmation test.[16] The employer who verifies the accuracy of a biochemical test greatly curtails the potential for biochemical discrimination. Second, after establishing the accuracy of the test, the employer should validate the biochemical test under the Uniform Guidelines on Employee Selection Procedures (the Uniform Guidelines) as promulgated by the Equal Employment Opportunity

10. Section 703(a)(2) of Title VII provides in pertinent part:

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

Id. § 703(a)(2), 42 U.S.C. § 2000e-2(a)(2).

11. Griggs v. Duke Power Co., 401 U.S. 424, 431 (1971) ("[Title VII] proscribes not only overt discrimination but also practices that are fair in form, but discriminatory in operation.").

12. *Id.*

13. *See infra* notes 79-103 and accompanying text.

14. *See infra* notes 67-78 and accompanying text.

15. *See infra* notes 79-112 and accompanying text.

16. *See infra* notes 74-78 and accompanying text.

Commission (EEOC).[17]

Part I of this note reviews the inherent inaccuracies of biochemical testing, suggesting that screening tests must be followed by confirmation tests.[18] Part II reviews the disparate impact theory of employment discrimination.[19] Part III suggests how biochemical testing might be implicated in disparate impact analysis.[20] Part IV discusses the different standards of validation under the two strands of disparate impact doctrine and distinguishes employee selection *tests* from employee selection *policies* that may be implemented through testing.[21] The section argues that courts must require employers to justify biochemical tests on the basis of their accuracy and job-relatednesss. Part V proposes a reproducible standard for validating biochemical tests using the Uniform Guidelines. The section concludes that employers must ensure the accuracy of biochemical tests by using high cutoff scores and confirmation tests.[22]

I. THE INHERENT INACCURACY OF BIOCHEMICAL TESTING

The validity of biochemical testing originates in the often fallible objectivity of scientific measurement. In order to consider more fully how a court should treat a biochemical test under a disparate impact analysis, this section first reviews the inherently inaccurate nature of biochemical testing.

A. *Screening Tests.*

Most biochemical tests begin with a screening test that is highly sensitive, but not highly specific.[23] Common screening tests for marijuana in the workplace are the radioimmunoassay (RIA) and the enzyme-multiplied immunoassay technique (EMIT).[24] A commonly used AIDS

17. *See* Equal Employment Opportunity Commission (EEOC) Uniform Guidelines on Employee Selection Procedures (1978), 29 C.F.R. § 1607.1(B) (1987).

18. *See infra* notes 23-38 and accompanying text.

19. *See infra* notes 39-66 and accompanying text.

20. *See infra* notes 67-78 and accompanying text.

21. *See infra* notes 79-112 and accompanying text.

22. *See infra* notes 113-30 and accompanying text.

23. *See* DiGregorio & Sterling, *Marijuana Pharmacology and Urine Testing,* 35 CLINICAL PHARMACOLOGY 209, 211-12 (1987); Evans, *Legal Issues in Alcohol and Drug Detection Programs,* 4 SYVA MONITOR, Spring 1986, at 1, 3; Gold & Dackis, *Role of the Laboratory in the Evaluation of Suspected Drug Abuse,* 47 J. CLINICAL PSYCHIATRY, Jan. 1986, supp. 1, at 17, 21-22; Morgan, *Problems of Mass Urine Screenings for Misused Drugs,* 16 J. PSYCHOACTIVE DRUGS 305, 305-06 (1984). For a discussion of AIDS screening tests, see Barry, Cleary & Fineberg, *Screening for HIV Infection: Risks, Benefits, and the Burden of Proof,* 14 LAW MED. & HEALTH CARE 259, 260-64 (1986).

24. *See* DiGregorio & Sterling, *supra* note 23, at 211-12; Gold & Dackis, *supra* note 23, at 19, 21.

screening test is the enzyme-linked immunosorbent assay (ELISA).[25]

One problem with screening tests is that they are *inferential*: they only detect the chemical by-products, or metabolites, of drugs or the AIDS virus, not the drug or virus itself.[26] Other substances have the same molecular shape as the by-products of marijuana and the AIDS virus; screening tests cannot differentiate between these "cross-reactive" impostors and the by-products of the AIDS virus or drugs. Cross-reactive substances fit the molecular "keyhole" of a biochemical test and cause a "false positive."[27] For example, one form of the EMIT will falsely register positive for marijuana when a common pain reliever, ibuprofen, is present in urine.[28]

The high sensitivity of marijuana screening tests is also problematic. A urine test detects minute quantities of marijuana by-products in the urine. Whether the amount detected represents a "positive" result depends solely on the arbitrary setting of a "cutoff level."[29] When metabolites detected are in excess of this predetermined cutoff level the result is

25. *See* Barry, Cleary & Fineberg, *supra* note 23, at 260.

26. *See id.*; Willette, *Interpreting Cannabinoid Assay Results,* 4 SYVA MONITOR, Winter 1986, at 1, 1-2; Gold & Dackis, *supra* note 23, at 20-21; Morgan, *supra* note 23, at 306, 309.

27. Barry, Cleary & Fineberg, *supra* note 23, at 20-21. Cross-reactions occur when compounds are chemically similar to drugs the test is designed to detect. Screening methods such as RIA and EMIT test for the presence of marijuana through an inferential chain of roughly five steps: 1) a urine sample is added to a cloudy bacterial suspension containing a drug-enzyme complex and a drug antibody; 2) the drug antibody binds to the drug-enzyme complex and inactivates the enzyme; 3) if the drug is also present in the urine sample, however, some of the antibody will bind to that free drug, rather than to the drug-enzyme complex, and the enzyme will remain active; 4) the active enzyme dissolves bacterial cell walls, and causes the cloudy bacterial suspension to become clear; 5) measuring the turbity of solution enables detection of the drug in the urine sample. *See* Morgan, *supra* note 23, at 306-08.

28. SYVA recently informed its client laboratories that the EMIT d.a.u. and EMIT st Cannabinoid 20 ng Assay, which use the enzyme malate dehydrogenase (MDH), could cross-react with the anti-inflammatory drugs ibuprofen (used in such nonprescription drugs as Motrin), naproxen (Naprosyn), and fenoprofen (Nalfon). Letters from SYVA to client laboratories (Feb., Mar. & July, 1986). In other words, a person taking Motrin, Naprosyn or Nalfon might have a positive EMIT urine test for marijuana even though the person had not been exposed to marijuana.

29. *See* DiGregorio & Sterling, *supra* note 23, at 212; Willette, *supra* note 26, at 2. Depending on the selected cutoff levels, EMIT may register positive days after marijuana nse. *Id.* at 2. If a cutoff level of 100 ng/ml is used, it is unlikely that a person would remain positive in the EMIT d.a.u. or EMIT st for more than three days after using a moderate amount of marijuana. If the cutoff is lowered to 20 ng/ml, the EMIT d.a.u. results may remain positive for two to seven days after the smoking of a single marijuana cigarette (containing about 120 mg of tetrahydrocannabinol (THC)). *Id.* (citing Bastiani, *Urinary Cannabinoid Excretion Patterns,* in THE CANNABINOIDS: CHEMICAL, PHARMACOLOGIC, AND THERAPEUTIC ASPECTS 263-80 (S. Agurell, W. Dewey & R. Willette eds. 1984)). Other studies report positive results with 20 ng/ml EMIT d.a.u. from 4 to 47 days (with occasional positive results for an additional period of up to 77 days in the case of heavy (daily) marijuana users). *Id.* at 3 (citing Ellis, Mann, Judson, Schramm & Tashchian, *Excretion Patterns of Cannabinoid Metabolites After Last Use in a Group of Chronic Users,* 38 CLINICAL PHARMACOLOGY & THERAPEUTICS 572 (1985)).

considered positive; a very low cutoff level may yield a positive result for marijuana over a month after ingestion.[30] Moreover, an EMIT test with a low cutoff level may indicate a positive result when an employee has had only passive exposure to marijuana; an employee could test positive by unintentionally inhaling the marijuana smoke of those around her,[31] just as non-smokers passively inhale the cigarette smoke of others. Further, due to the fat-soluble nature of marijuana metabolites, urine screening tests may continue to show a positive result when metabolites leach out of fat cells long after actual intoxication.[32] A marijuana screening test, therefore, cannot measure impairment; it can only inferentially measure ingestion.

Thus, courts and employers cannot rely on screening tests alone as an indicator of drug use that might compromise job performance. Because of cross-reactivity and oversensitivity, positive results in a screening test will often have no relationship to impairment or, more importantly, job performance. One way to improve the accuracy of screening tests is to raise the cutoff level, but the problem of cross-reac-

30. *See* DiGregorio & Sterling, *supra* note 23, at 210; Willette, *supra* note 26, at 1. A major problem with urine tests is that they do not distinguish the presence of drugs that have been ingested recently from drugs that have been present in the body for a longer period of time. This problem is especially severe in marijuana testing. *See id.*, at 1. THC, the psychoactive ingredient in marijuana, is a soluble drug and may be stored in the fat cells and later released over a period of several days or weeks. Screening tests for THC remain positive while THC leaches out of the fat cells. *See id.* This phenomenon is particularly striking for chronic users of marijuana, who have large amounts of THC stored in their fat cells. One chronic user had positive test results eleven weeks after discontinuing all use of marijuana. *Id.* The typical screening test for non-chronic users will remain positive for one to three days. Cohen, *supra* note 6, at 7.

31. *See* Cone, Johnson, Darwin, Yousefnejad, Mell, Paul & Mitchell, *Passive Inhalation of Marajuana Smoke: Urinalysis and Room Air Levels of Delta-9-Tetrahydrocannabinol,* 11 J. ANALYTICAL TOXICOLOGY, 89, 95-96 (1987) (concluding that passive inhalation of marijuana smoke can lead to excretion of detectable levels of cannabinoids in urine); Cone & Johnson, *Contact Highs and Urinary Cannabinoid Excretion After Passive Exposure to Marijuana Smoke,* 40 CLINICAL PHARMACOLOGY & THERAPEUTICS 247, 255 (1986) (Passive exposure to 16 marijuana cigarettes for one hour each day for six days produced urine drug levels ranging from 20 ng/ml to as much as 100 ng/ml; cutoff level for a positive result was 20 ng/ml.); *see also* Mason, Perez-Reyes, McBay & Foltz, *Cannabinoids in Plasma After Passive Inhalation of Marijuana Smoke,* 249 J. A.M.A. 475, 476 (1986) (letter to editor describing findings of low level THC concentrations in passive exposed subjects, but noting distinguishable differences between active and passive exposures). *But see* Willette, *supra* note 26, at 3-5 (deemphasizing the importance of passive inhalation).

32. *See* DiGregorio & Sterling, *supra* note 23, at 212; Willette, *supra* note 26, at 2-5. John P. Morgan, M.D., states that the presence of marijuana in the urine does not correlate with actual intoxication. Marijuana persists in the blood and urine long after the period of intoxication ends. Morgan, *supra* note 23, at 306. Measuring actual intoxication is important since a criterion-related validity study would need to correlate positive test results with measurable criteria such as error rate. Some experts have proposed that one way to improve measurement of intoxication is by monitoring brain waves. Herzfeld, *Brain Scans on the Job,* AM. HEALTH, Jul.-Aug. 1986, at 72, 74-76. A preliminary brain wave test has been designed. *Id.*

tivity remains. Only confirmation tests come close to eliminating both oversensitivity and cross-reactivity.

B. *Confirmation Tests.*

Confirmation tests eliminate much of the inaccuracy in biochemical testing and are more specific than screening tests.[33] They tend to locate false positives and confirm the initial readings of a screening test only when a true positive is present. The most effective confirmation test for drug testing is gas chromatography-mass spectrometry (GC-MS), a test so precise that it is often called "molecular fingerprinting."[34] The most common test to confirm an initial positive for the AIDS virus is the Western blot assay.[35]

---

33. In other words, they tend to register positive only when the biochemical measured is present. For example, the generally accepted process used to detect marijuana is to administer a screening test. Positive results may be confirmed with the more specific gas chromatography-mass spectrometry (GC-MS) test. Evans, *supra* note 23, at 3; Gold & Dackis, *supra* note 23, at 21-22.

34. Gold & Dackis, *supra* note 23, at 21. Gas chromatography (GC) employs an expensive instrument called a chromatographer. At a given temperature and pressure, *identical compounds,* such as molecules of a particular drug, will travel through a spiral column at the same speed, since their interaction with the column packing is the same. Molecules of *different* compounds such as marijuana and cocaine, by contrast, will travel down the column at different speeds, thereby allowing differentiation and classification. *Id.*

The "sifting" process occurs in the chromatographer, while the classification process occurs in the mass-spectrometer (MS). After passing through the GC, the molecules are subjected to a barrage of high-energy particles designed to fragment the marijuana molecule. A marijuana molecule will break apart in a characteristic way because not all of the bonds holding its atoms together have the same strength. *See id.* The presence of a drug in the urine is confirmed if the fragmentation pattern of a molecule in the sample matches that of the drug sought to be detected. *Id.*; Interview with Robert G. Kaley, Ph.D., Analytical Chemist, Monsanto Chemicals, Inc. (Aug. 4, 1986) (discussing high specificity and sensitivity of GC-MS).

The GC-MS is the preferred test for confirming the presence of marijuana or marijuana metabolites. *See* Jones v. McKenzie, 628 F. Supp. 1500, 1503 (D.D.C. 1986); Higgs v. Wilson, 616 F. Supp. 226, 229 (W.D. Ky. 1985). Notably, however, even the GC-MS test is based on biochemical inference rather than observation.

35. Barry, Cleary & Fineberg, *supra* note 23, at 260 (Western blot assay is not practical for large-scale screening because it is labor-intensive; it is used to confirm positive antibody tests designed for mass screening.); *see also* Essex, Allen, Kanki, McLane, Malone, Kitchen & Lee, *Antigens of Human T-Lymphotropic Virus Type III/Lymphadenopathy-Associated Virus,* 103 ANNALS INTERNAL MED. 700, 701 (1985) (Western blot test, under certain circumstances, "can be a valuable confirmatory assay."); Gershoni & Palade, *Protein Blotting: Principles and Applications,* 131 ANALYTICAL BIOCHEM. 1, 3 (1983) (explaining the protein blotting process).

In the Western blot assay, HIV proteins grown in a tissue culture are separated into component proteins and "blotted" onto special paper. A blood sample is added, and any antibody present in the blood is "sandwiched" by a radioactive antibody probe. Antibody complexes are detected either by exposing the paper to X-ray film, creating "hotspots," or by adding the enzyme's substrate to the antibody complex, causing a color reaction. The Western blot test also generates "hotspots" concentrated in bands, allowing more precise detection of specific HIV proteins. The Western blot assay, however, is more expensive and more labor-intensive than the ELISA mass-screening test. Barry, Cleary & Fineberg, *supra* note 23, at 260.

Employers, however, do not always use confirmation tests because they usually are more expensive and more labor-intensive than screening tests.[36] Nonetheless, use of confirmation testing can only benefit an employer. A truly job-related, nondiscriminatory test will deter frivolous litigation and contribute to safety, efficiency, and profit maximization in the workplace.[37] Many courts have required that screening tests be followed by confirmation tests.[38]

II. THE DISPARATE IMPACT THEORY OF EMPLOYMENT DISCRIMINATION

Congress intended that Title VII eliminate employment discrimination on the basis of race, sex, religion, or national origin.[39] In the seminal case of *Griggs v. Duke Power Co.*,[40] the Supreme Court set forth a theory of employment discrimination under Title VII known as "disparate impact."[41] The disparate impact theory seeks to remedy the effect of past discrimination and to remove "artificial, arbitrary, and unnecessary barriers to employment."[42]

Since the *Griggs* decision was handed down, however, disparate impact doctrine has been less than coherent. Issues as basic as what weight a court should give the Uniform Guidelines remain unresolved.[43] Moreover, lower courts recognize a variety of methods for establishing a prima facie claim of disparate impact.[44] More important, *Griggs* itself uses several different phrases to describe the standard of an employer's business justification.[45] The most recent point of contention involves the question of whether the scope of disparate impact testing includes subjective hir-

36. *See* Rust, *Drug Testing: The Legal Dilemma,* A.B.A. J., Nov. 1, 1986, at 50, 52 (One company charges from $100 to $200 per specimen for a confirmatory GC-MS test.).

37. *See infra* notes 74-78 and accompanying text.

38. *See* Jones v. McKenzie, 628 F. Supp. at 1507 ("[T]ermination of plaintiff's employment on the basis of an unconfirmed EMIT test was arbitrary and capricious."); Higgs v. Wilson, 616 F. Supp. at 232 (inmates entitled to preliminary injunction against disciplinary actions by corrections officials based on unconfirmed EMIT test for marijuana); *see also* Storms v. Coughlin, 600 F. Supp. 1214, 1221-22 (S.D.N.Y. 1984) (Unconfirmed EMIT urine test results for marijuana are unreliable enough to warrant challenge to drug testing program designed by New York prison officials.)

39. *See supra* note 10 and accompanying text (language of statute at 42 U.S.C. §§ 2000e to 2000e-17 (1982)).

40. 401 U.S. 424 (1971).

41. *Id.* at 429-36. "The Act proscribes . . . practices that are fair in form, but discriminatory in operation." *Id.* at 431.

42. *Id.* at 431-32.

43. *See* 3 A. LARSON, EMPLOYMENT DISCRIMINATION §§ 75.60-.65 (1987).

44. *See infra* note 57 and accompanying text.

45. *See infra* notes 59-63 and accompanying text.

ing practices.[46]

In *Griggs,* the Supreme Court ruled that employment practices having a disparate impact on a protected class violate the prohibitions of Title VII unless the employer can show that the selection device is "job-related."[47] Before the effective date of Title VII in 1965, Duke Power hired blacks only to work in its labor department.[48] After the enactment of Title VII, Duke Power abandoned its policy of overt discrimination and substituted testing procedures. The company required that applicants to departments other than labor have a high school diploma or that they achieve satisfactory scores on written aptitude tests.[49] The plaintiffs argued that although these requirements applied equally to blacks and whites, they disproportionately excluded blacks from hiring and promotion.[50] The Court found that Duke Power was unable to justify the employment practices.[51]

Griggs and its progeny[52] set forth the core of current disparate impact doctrine. There are two steps that must be satisfied. First, a plaintiff must make a prima facie showing that a facially neutral selection procedure has a disparate impact on his or her protected group.[53] To establish a *significant* discriminatory effect, plaintiffs usually produce statistical data. For example, the *Griggs* Court relied on evidence that 12% of black males and 34% of white males held high school diplomas in North Carolina, and on an unrelated EEOC finding that only 6% of

46. *See* Watson v. Fort Worth Bank & Trust, 798 F.2d 791, 808 n.17 (5th Cir. 1986) (Goldberg, J., dissenting), *vacated and remanded,* 108 S. Ct. 2777 (1988).

47. 401 U.S. 424, 431 (1971).

48. *Id.* at 427. Duke Power organized its Dan River, N.C. plant into five departments: (1) labor, (2) coal handling, (3) operations, (4) maintenance, and (5) laboratory and test. Black employees brought the action in *Griggs* because the highest paying jobs in the labor department paid less than the lowest paying jobs in the other four "operating" departments. *Id.*

49. *Id.* at 427-28. The tests were the Wonderlic Personnel Test, which measures intelligence, and the Bennett Mechanical Comprehension Test. *Id.* at 428.

50. *Id.* at 429-30. Rejecting the denial of the claims by lower courts, Chief Justice Burger reasoned that lower scores on intelligence tests were traceable to the effects of past discrimination, such as inferior education and segregation. *Id.* at 430.

51. *Id.* at 431-32. "What is required by Congress is the removal of artificial, arbitrary, and unnecessary barriers to employment when the barriers operate invidiously to discriminate on the basis of racial or other impermissible classification." *Id.* at 431. An employer could justify the practice only if it was a "business necessity" or was "related to job performance." *Id.* Good intent could not exonerate a practice if it operated as a "built-in headwind[ ]." *Id.* at 432.

52. *See, e.g.,* Connecticut v. Teal, 457 U.S. 440, 452-56 (1981) (holding that a higher percentage of black candidates being promoted than white candidates does not preclude blacks from making prima facie case of disparate impact nor does it provide defendant with a defense); Albemarle Paper Co. v. Moody, 422 U.S. 405, 425 (1975) (even if job-relatedness shown, plaintiff may prevail by demonstrating test is a "pretext" for discrimination).

53. *See* 401 U.S. at 430 ("Under the Act, practices, procedures, or tests neutral on their face, and even neutral in terms of intent, cannot be maintained if they operate to 'freeze' the status quo of prior discriminatory employment practices.").

blacks, but 58% of whites, had passed the disputed tests.[54]

The statistical approach used in *Griggs* is instructive but not binding. Although the Supreme Court extensively discussed the many issues associated with statistical evidence of employment discrimination in *Hazelwood School District v. United States,*[55] it has yet to endorse a particular statistical method for showing a prima facie claim of disparate impact.[56] The lower federal courts recognize several methods of statistical proof, most notably the two-tailed test of standard deviation and the 80% rule.[57]

Once a plaintiff has established a prima facie claim of disparate impact, the second step of the *Griggs* analysis follows. In this step, the defendant assumes the burden of justifying the employment practice in question.[58] What justifies an employment practice is not entirely clear, although *Griggs* appears to require that an employer justify an employment practice by showing that it is "job-related."[59] In subsequent decisions, however, the Court has emphasized the *Griggs* language requiring an employer to show the "business necessity" of an employment practice.[60] In cases not involving objective tests, the Court has not required

54. *Id.* at 430 & n.6.

55. 433 U.S. 299, 307-13 (1977) (holding that the proper statistical comparison was between the racial composition of the school district's teaching staff and the racial composition of the qualified public school teacher population in the relevant labor market, not the racial composition of the student population).

56. *See generally* 3 A. LARSON, *supra* note 43, § 74 (discussion of methods of proof).

57. *See* Fudge v. City of Providence Fire Dep't, 766 F.2d 650, 658 nn.8-9 (1st Cir. 1985) (50% statistical significance test); Easley v. Anheuser-Busch, Inc., 758 F.2d 251, 256 n.8 (8th Cir. 1985) (80% test); Firefighters Inst. for Racial Equality v. City of St. Louis, 616 F.2d 350, 356-57 (8th Cir. 1980) (80% test); *see also* G. KIMBLE, HOW TO USE (AND MISUSE) STATISTICS 116-18 (1978) (explaining why standard deviation is a representative measure of variability); Meier, Sacks & Zabell, *What Happened in* Hazelwood: *Statistics, Employment Discrimination, and the 80% Rule,* 1984 AM. B. FOUND. RES. J. 139, 167-68 (1984) (arguing that the 80% rule should be favored over the imprecise *Hazelwood* rule); Kaye, *The Numbers Game: Statistical Inference in Discrimination Cases,* 80 MICH. L. REV. 833, 839-41 (1982) (reviewing D. BALDUS & J. COLE, STATISTICAL PROOF OF DISCRIMINATION (1980) and noting fallacies in *Hazelwood* approach).

58. 401 U.S. at 432.

59. *Id.* at 431.

60. *Id.* at 433-36 & n.9 (Court cited EEOC Guidelines on Employment Testing Procedures and treated them "as expressing the will of Congress."). The EEOC issued its first employment guidelines in 1966 as the Guidelines on Employment Testing Procedures. The first procedures were not codified, however. In 1970, the EEOC published the EEOC Guidelines on Employee Selection that superseded and enlarged the 1966 guidelines. 29 C.F.R. § 1607 (1971). "The guidelines . . . are based on the belief that properly validated and standardized employee selection criteria can significantly contribute to the implementation of non-discriminatory personnel policies, as required by Title VII." *Id.* § 1607.1. The regulations detailed testing and validation procedures that differ significantly from the present guidelines in their emphasis on criterion-related validation. The original guidelines also prohibited use of validated tests in two circumstances on which the new Uniform Guidelines are much less rigorous. *See* 3 A. LARSON, *supra* note 43, § 75.52.

strict validation at all.[61]

The *Griggs* Court accorded "great deference" to the old EEOC Guidelines as a method for validating employee selection tests;[62] Duke Power was required to "validate" the test under a method based primarily on the old EEOC Guidelines.[63] More recently, the Supreme Court has retreated from strict reliance on the EEOC Guidelines.[64] The Guidelines as they existed when the Court decided *Albemarle Paper Co. v. Moody* imposed a harsh burden on employers;[65] the Supreme Court and lower federal courts have explored other, more workable validation methods for tests and for non-test devices, typically relying indirectly on

---

Some employers were subject to both the EEOC Guidelines and the Office of Federal Contract Compliance (OFCC) Guidelines, 60 C.F.R. § 60-3 (1977), which were similar to the old EEOC Guidelines on a number of points, differing primarily on the issue of presentation of the evidence of validation and sanctions available for violation of the OFCC Guidelines. *See* 3 A. LARSON, *supra* note 43, § 75.52, at 15-26 to -27. The OFCC regulations applied to contractors and subcontractors of the U.S. government and were administered by the Department of Labor, 41 C.F.R. § 60-3 (1977), while the EEOC Guidelines applied to employers covered by Title VII of the Civil Rights Act of 1964 and were administered by the EEOC, 29 C.F.R. § 1607.1(c). *See* EEOC, COORDINATION OF FEDERAL EQUAL EMPLOYMENT OPPORTUNITY PROGRAMS: THE FIRST YEAR 1978-1979, at 3 (Aug. 1979). The resulting conflict between the two sets of regulations led to the formation in 1972 of the Eqnal Employment Opportuuity Coordination Council. *Id.* After several years the Council, whose functions were eventually tranferred to the EEOC, successfully consolidated the two sets of testing guidelines into the Uniform Guidelines on Employee Selection Procednres. 29 C.F.R. § 1607 (1979). The Uniform Guidelines essentially require employers to examine their selection processes for adverse impacts of discrimination and, if adverse impacts are found, to either: (1) eliminate the adverse impact; (2) validate the impact and search for new alternatives; or (3) otherwise justify the procednre under federal law. *See* EEOC, *supra,* at 3; *see also* Adoption of Questions and Answers to Clarify and Provide a Common Interpretation of the Uniform Guidelines on Employee Selection Procedures, 44 Fed. Reg. 11,996 (1979).

61. *See, e.g.,* Dothard v. Rawlinson, 433 U.S. 321, 331 n.14 (1977) (To survive Title VII challenge, discriminatory employment practice must be "necessary to safe and efficient job performance."). This case really involved a *measurement* of height and weight; perhaps this factor explains why the Court refused to allow the low standards of business justification entailed in the non-testing strand.

62. *See, e.g.,* Beazer v. New York City Transit Auth., 440 U.S. at 587 n.31 (finding that the legitimate goals of safety and efficiency justified defendant's drug-related hiring policy).

63. 401 U.S. at 433 n.9 (interpreting professional nature of "professionally developed ability test").

64. *See, e.g.,* Washington v. Davis, 426 U.S. 229, 247 n.13 (1976) ("It appears beyond doubt by now that there is no single method for appropriately validating employment tests for their relationship to job performance."); Connecticut v, Teal, 457 U.S. 440, 452-53 & n.12 (1981) (rejecting bottom-line defense of Guidelines after addressing issue of "job-related tests," but failing to mention validation).

65. *See, e.g.,* Albemarle Paper Co. v. Moody, 422 U.S. 405, 431-36 (1975) (setting rigid standards for validating test while claiming "appropriate standard of proof for job relatedness has not been clarified until today"); Douglas v. Hampton, 512 F.2d 976, 986 (D.C. Cir. 1975) (conrt will consider construct validity only after a showing that proof of criterion validity is infeasible). *See generally* Vulcan Soc'y v. Civil Serv. Comm'n, 490 F.2d 387, 394-96 n.10 (2d Cir. 1973) (discussing difficulties of strict validation under *Albemarle*).

the Guidelines.[66] It is important to note, however, that even though the method of validation fluxes, the requirement that validation occur does not.

III. BIOCHEMICAL DISCRIMINATION

Under the disparate impact theory a plaintiff need not prove why a test is discrminatory. To establish a prima facie case, a plaintiff must only make the statistical showing required in that particular jurisdiction. As this note discusses below, the disparate impact analysis applies to biochemical tests.[67] Thus, when an employer uses a biochemical test as a pass/fail barrier to employment, a plaintiff need only show that the test excludes a disproportional number of a protected class, not how the test excludes them. Once a plaintiff makes this showing, the employer must articulate a business justification for the biochemical test. Because of a biochemical test's potential inaccuracy, however, an employer must also prove the accuracy of a biochemical test.

A biochemical test might have a disparate impact on the members of a protected class for a number of reasons, but the most troubling is that the test itself may subtly discriminate regardless of the employment policy it enforces. For example, it is possible that some substance peculiar to the physiology of a given race may be cross-reactive and thereby cause false positives in common screening tests like the RIA or EMIT.[68] Ac-

66. Either content, construct or criterion methods of validation are available as alternative approaches under the Guidelines. *Davis,* 426 U.S. at 247 n.13. The circuit courts have accepted the Guidelines as expert advice, but not as binding authority. *See, e.g.,* Conteras v. City of Los Angeles, 656 F.2d 1267, 1281 (9th Cir.), *cert. denied,* 455 U.S. 1021 (1982); Guardians Ass'n v. Civil Serv. Comm'n, 630 F.2d 79, 90-91 (2d Cir. 1980), *cert. denied,* 452 U.S. 940 (1981); *see also* Thompson & Christiansen, *Court Acceptance of Uniform Guidelines Provision: The Bottom Line and the Search for Alternatives,* 8 EMPLOYEE REL. L.J. 587, 598-602 (1983); Booth & Mackay, *Legal Constraints on Employment Testing and Evolving Trends in the Law,* 29 EMORY L.J. 121, 141, 164-65 (1980).

The key point is that, while the courts need not accept the Guidelines as a standard for validating biotests, they should because such acceptance would promote the efficiency, predictability and equity that comes from rigid but reproducible scientific and professional rules.

67. *See infra* notes 68-73 and accompanying text.

68. Dr. James Woodford, a forensic chemist, advanced a "melanin theory" of cross-reactivity. The theory was used by fired black police cadets to challenge the accuracy of urine drug testing by police officials. Lindsey, *Worker Drug Test Provoking Debate,* N.Y. Times, May 3, 1986, at 1, col. 3. According to this theory, melanin, a dark skin pigment, can break down into a substance that causes false positives in marijuana urine tests, increasing the risk of racial bias. *Id.* The theory inspired the ACLU to bring a case in Cleveland. Shield Club v. City of Cleveland, 647 F. Supp. 274, 285 (N.D. Ohio 1986) (finding Woodford's theory to be unreliable without further "creditable and authoritative data."), *rev'd mem.,* 834 F.2d 172 (6th Cir. 1987).

Woodford's theory is incorrect on several grounds. First, although melanin is dispersed differently in blacks and whites, blacks and whites have about the same number of melanin cells or melanocytes. Warshauer & Steinbaugh, *Sunlight and Protection of the Skin,* 27 AM. FAM. PHYSICIAN, June 1983, at 109, 110. Second, even if there were a difference in the number of melanin cells, the

cording to one study, low-risk populations—such as caucasian women and Asians—might have a false positive rate as high as 88.7%.[69] The study theorizes that the Western blot confirmation tests will be accurate in 71.8% of cases.[70]

Biochemical tests may also reflect socioeconomic patterns of drug use and AIDS infection. It is well-documented that AIDS is more preva-

---

lining of the urinary tract would contain almost no melanin. *See* GRAY'S ANATOMY 43 (36th ed. 1980). Third, a study conducted by Roche Diagnostic Systems, makers of the RIA test, indicated no interference or cross-reactivity by melanin or its metabolite 5-hydroxyindole. RADIOIMMUNOASSAY TRAINING PROGRAM FOR DRUGS OF ABUSE TESTING ON THE RELATIVE REACTIVITY, SPECIFICITY AND CROSS REACTIVITY OF THE ABUSCREEN REAGENT SYSTEM (June 30, 1985). Fourth, no interference was found in one published study under the GC-MS test. ElSohly, Jones, ElSohly & Stanford, *Analysis of the Major Metabolite of Delta-9-tetrahydrocannabinol in Urine: VI. Specificity of the Assay with Respect to Indole Carboxylic Acids,* 9 J. ANALYTICAL TOXICOLOGY 190, 191 (1985).

While Woodford's theory is suspect, it is possible that some other ethnically characteristic substances might cause one ethinic group to test falsely positive more than another. For example, blacks and Hispanics might have a greater tendency to test positive because they excrete higher (or perhaps different) levels of MDH and lysozyme in their urine. *See* Morgan, *supra* note 23, at 312.

Indeed, several molecules have ethnically idiosyncratic patterns of occurrence, accounting for cetain well-known genetic diseases. American blacks characteristically have a 1:600 chance of acquiring sickle cell anemia, with a 1:12 chance of having the partial defect, the sickle cell trait. Lubin & Mentzen, *Sickle Cell Disease,* in PEDIATRICS 1068, 1072 (A. Randolph 17th ed. 1982).

Another ethnically characteristic biochemical alteration is thalassemia, also due to a single gene deletion. *Id.* at 1074. Alpha-Thalassemia is most common in Southeast Asians, but "is found in 2 to 7 percent of American black newborns." *Id.* B-Thalassemia occurs in more than five percent of Italians, Greeks, Sardinians, Sicilians, Indians and Southeast Asians. *Id.* Thalassemia may result in such complications as severe anemia, growth retardation, organ damage, and fatal cardiac failure. *Id.* at 1075. By contrast, the Tay-Sachs disease—also due to an enzyme deficiency—strikes Ashkenazi Jews and is characterized by multisystem defects. Brady, *Sphingolipidoses,* in PEDIATRICS, *supra,* at 311, 316-317.

Over 100 million people throughout the world suffer from the G6PD deficiency. Lubin & Mentzer, *Abnormalities of Erythrocyte Metabolism,* in PEDIATRICS, *supra,* at 1082, 1083. Ironically, the EMIT d.a.u. assay for marijuana uses the glucose 6 phosphate dehydrogenate enzyme (G6PDH) in its inferential chain. EMIT Cannabinoid User's Guide (Feb. 1986); letter from SYVA to laboratories (July 1986). G6PD deficiency is uncommon in non-Mediterrancan Caucasians. Lubin & Mentzer, *supra,* at 1083. The deficiency affects over one percent of Mediterranean and Middle Eastern males. It affects five percent of Chinese males and ten percent of American black males. *Id.* While there is no scientifically established link between false positives in biochemical tests and these enzyme deficiencies, it would be advisable to keep factors such as these in mind when evaluating future biotests for their discriminatory potential. These dangers of discrimination underscore the need for strict *professional* standards of validations, in accord with section 703(h).

69. *See* Barry, Cleary & Fineberg, *supra* note 23, at 263. These differences trace to the fact that differing prevalences affect *predictive value.* The false-positive *rate* might be the same in the low-risk and the high-risk population. But because more people would be uninfected in the low-risk population, a greater percentage of the positives would be false positive than in the high-risk population. *See* C. PIERCE & D. VAN DEVEER, AIDS: ETHICS AND PUBLIC POLICY 142-44 (1988) (describing predictive value, preference and false positives); Ransohoff & Feinstein, *Problems of Spectrum and Bias in Evaluating the Efficacy of Diagnostic,* 299 NEW ENG. J. MED. 929, 926-27 (1978) (same).

70. Hoyt, Finnigan, Nee, Shults & Butler, *Drug Testing in the Workplace—Are Methods Legally Defensible?,* 258 J. A.M.A. 504, 507 (1987).

lent among blacks and hispanics than among other racial groups.[71] Because AIDS strikes certain groups with greater frequency, an AIDS screening test might be discriminatory in effect if what the test purports to measure is not job-related.[72] Also, in certain locales, marijuana use rates among a minority group may exceed the rates of other groups. Non-user minorities in these locales have a greater likelihood of passive exposure to marijuana smoke.[73] These examples are only possible explanations for a statistical instance of disparate impact. Under *Griggs*, however, the cause of the disparity would be irrelevant.

Confirmation tests greatly reduce the potential for biochemical discrimination present at the screening test level. The GC-MS test for marijuana effectively eliminates most cross-reactive substances.[74] In addition, if the GC-MS test is preceded by a screening test that raises the cutoff level for positive results to 100 nanograms per milliliter, the probability of detecting passive exposure or off-duty use is minimized.[75] Likewise, if cutoff levels for the ELISA test for AIDS are raised in low-risk populations and the ELISA test is confirmed with the Western blot assay, the numbers of false positives decrease.[76] Even if an AIDS biochemical test is positive, however, an employer would need to prove that a positive test result is job-related. Such proof would be difficult unless the patient were afflicted with debilitating AIDS encephalopathy or brain disease.[77] A positive AIDS test without such accompanying manifestations would probably not be sufficiently job-related since the majority of research in-

71. A CDC report emphasizes the disproportionate occurrence of AIDS in blacks and Hispanics, especially women and children. CDC, *Acquired Immunodeficiency Syndrome (AIDS) Among Black and Hispanics—United States*, 35 MORBIDITY & MORTALITY WEEKLY REP. 655, 655-66 (1986); *see also* Mueller, *The Epidemiology of the Human Immunodeficiency Virus Infection*, 14 LAW MED. & HEALTH CARE 250, 252 (1986). Likewise, a survey involving screening for military recruits replicated these results, showing the following rates: blacks (3.9/1000), other nonwhites (2.6/1000) and whites (0.9/1000). CDC, *Human T-Lynphotropic Virus Type III/Lymphadenopathy-Associated Virus Antibody Prevalence in U.S. Military Recruit Applicants*, 35 MORBIDITY & MORTALITY WEEKLY REP. 421, 421-24 (1986). These later prevalence rates are significant because they show that over four times as many black military recruits are infected as white recruits (0.9 x 4 = 3.6).

72. See *supra* note 5 and accompanying text for a discussion of the symptoms of AIDS that might and might not affect an employee's job performance.

73. *See* 1 (4) NIDA STATISTICAL SERIES: ANNUAL DATA 70-104 (1984) (Federal data suggest that marijuana use rates are related to socioeconomic status and there will be heavier concentrations of use in low to lower middle classes. In certain urban areas, minority use rates will far exceed nonminority rates, but the reverse will be true in other cities.).

74. *See* DiGregorio & Sterling, *supra* note 23, at 211; Gold & Dackis, *supra* note 23, at 20-22.

75. *See supra* note 29 and accompanying text.

76. *See* DiGregorio & Sterling, *supra* note 23, at 210; Willette, *supra* note 26, at 2.

77. *See supra* note 5 and accompanying text.

dicates that the disease cannot be transmitted by casual contact.[78] Only in jobs involving health care might an AIDS test be job-related, although an employer could make an argument that AIDS testing is job-related because of the potential rise in health insurance premiums.

IV. Biochemical Tests as Employee Selection Devices

Griggs and its progeny establish two strands of disparate impact doctrine: the selection policy strand and the professionally developed testing strand. This section discusses both strands and explains why biochemical tests comfortably fit within traditional disparate impact definitions of employee selection tests. Finally, this section distinguishes a test from the policy it implements. When the accuracy of the test itself is called into question, an employer seeking to prove its job-relatedness must offer evidence of the test's accuracy *and* validate the test.

A. *The Selection Policy Strand.*

In several decisions, the Supreme Court has applied the disparate impact doctrine to non-test selection devices. These decisions typically involve an employer's policy requiring employees to meet specified criteria or possess certain qualities. These decisions differ from testing decisions because the method of assessing or measuring the employer's requirement is not in question; it is the policy that the employer must justify. Consequently, in these decisions, the Court has not required validation under the Uniform Guidelines.

In *Dothard v. Rawlinson,*[79] the Court considered a non-test selection policy—Alabama statutes that established minimum height and weight requirements for applicants for the position of prison guard.[80] The requirements would have excluded 41.13% of American women while excluding less than 1% of men.[81] The Court did not require proof of the accuracy of a weight scale or of a yardstick; these were implicitly accepted as valid and accurate. The Court scrutinized the *policy* of requiring a minimum weight and height, not the methods of measurement. The employer argued that the policy selected applicants on the basis of strength.[82]

78. *See* S. Broder, AIDS: Modern Concepts and Therapautic Challenges 84 (1987); C. Pierce & D. Van DeVeer, *supra* note 69, at 3-5; Surgeon General's Report on Acquired Immune Deficiency Syndrome 1-36 (1986).

79. 433 U.S. 321 (1977).

80. *Id.* at 323-24.

81. *Id.* at 329-30.

82. *Id.* at 331.

The Court held that the employer had not shown the job-relatedness of the selection policy for two reasons. First, the Court reasoned that the employer could have used an alternative test that directly measured strength and defensive capacity.[83] Second, in a footnote, the Court stated, "A discriminatory employment practice must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge."[84] The Court did not require validation nnder the Uniform Guidelines, but the alternative test rationale is considered in the Guidelines.[85]

In *New York City Transit Authority v. Beazer,*[86] the Court also considered a non-test selection policy. In *Beazer,* the employer prohibited the hiring or promotion of workers using any narcotic,[87] including the legal therapeutic agent methadone, prescribed to treat heroin addiction.[88] When the employer discovered that employees were enrolled in methadone maintenance programs, the employees were dismissed.[89] The employer also used urine testing as a method of discovering drug abuse on the job.[90] The urine test was one mode of implementing what was primarily a selection *policy* rather than a pass/fail barrier to employment. The Transit Authority justified the blanket policy as necessary for the safe and efficient performance of work that conld potentially endanger the public.[91] The Supreme Court agreed that the "legitimate employment goals of safety and efficiency" justified the employment policy.[92]

The selection policy in *Beazer* implicated Title VII because the plaintiffs produced statistics showing that more blacks and hispanics than whites are involved in methadone maintenance programs.[93] The *Beazer* Court, however, rejected the statistical basis of the plaintiffs' claim.[94] Arguably, the plaintiffs did not establish a prima facie claim of

83. *Id.* at 332. The Court, referring the defendants to the EEOC Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607 (1976), said that a test that measures strength directly, "fairly administered, would fully satisfy the standards of Title VII because it would be one that 'measure[s] the person for the job and not the person in the abstract.' " 433 U.S. at 332 & n.15 (quoting Griggs v. Duke Power Co., 401 U.S. 425, 436 (1971)).

84. 433 U.S. at 330 n.14.

85. *See id.* at 332 & n.15 (giving the Court's only reference to the Guidelines); 29 C.F.R. § 1607 (1987).

86. 440 U.S. 568 (1979).

87. *Id.* at 571-72.

88. *Id.* at 573-75.

89. Beazer v. New York City Transit Auth., 399 F. Supp. 1032, 1036 (S.D.N.Y. 1975).

90. Positive urine tests for methadone, however, were not among the criteria used to identify suspected violators to the transit authority medical director. *Beazer,* 440 U.S. at 585 n.26.

91. *Beazer,* 399 F. Supp. at 1036.

92. *Beazer,* 440 U.S. at 587 n.31.

93. *Id.* at 579.

94. *Id.* at 584-87 (holding that the statistics presented "do not prove a violation of Title VII").

disparate impact, and the Court did not formally address the issue of proving the job-relatedness of the policy. If *Beazer* did not present a claim of disparate impact, the language the Court devoted to the job-relatedness issue[95] is dicta. In any event, the Court's dicta indicate a less demanding standard for proving job-relatedness in the non-test selection policy context.

Beazer may settle the question of whether the *policy* of excluding drug users from employment is related to the job. But *Beazer* does not address urine *testing*; the proper standard for proving the job-relatedness of a potentially inaccurate test is still validation under the Uniform Guidelines.

B. *The Testing Strand.*

Griggs and *Albemarle Paper v. Moody*[96] are the controlling decisions when the selection device is a test. Although in *Griggs* and *Albemarle* the Court relied on the EEOC Guidelines, in subsequent decisions the Court has shown less deference to the Guidelines.[97] Nevertheless, *Griggs* and *Albemarle* are still good law. Thus, while no longer entitled to great deference, the Uniform Guidelines continue to guide courts and employers when the selection device in question is a test.

Washington v. Davis[98] is the only Supreme Court decision involving a selection test that does not require a demanding standard of validation. In *Davis,* the District of Columbia police department required applicants to pass a written examination measuring verbal and communicative skills.[99] The police department argued that a positive correlation between test results and performance in a training program satisfied the job-relatedness requirement.[100] The Court found this justification sufficient to validate the test.[101] The *Davis* holding, however, is of limited value in the Title VII context. The Court grounded its decision on the

95. *Id.* at 587 ("At best, respondents' statistical showing is weak; even if it is capable of establishing a prima facie case of discrimination (and the rule's application to methadone users) it is assuredly rebutted by TA's demonstration that its narcotics rule is 'job related.' ").

96. 422 U.S. 405 (1975).

97. *See* Connecticut v. Teal, 457 U.S. 440, 453 n.12 (1982) (declining to defer to the Guidelines); General Elec. Co. v. Gilbert, 429 U.S. 125, 141-43 (1976) (applying the Stridmore v. Swift, 323 U.S. 134, 140 (1944), standard on the role of interpretative rulings); Washington v. Davis, 426 U.S. 229, 263 (1976) (Brennan, J., dissenting).

98. *Id.* at 229.

99. *Id.* at 234-35.

100. *Id.* at 235-36.

101. *Id.* at 250-52.

plaintiffs' constitutional claim and not Title VII.[102] Most lower courts have limited the application of *Davis,* requiring more than correlation between test scores and success in training for validation.[103]

C. *Biochemical Tests as Selection Tests.*

Disparate impact decisions of the testing strand have typically involved paper-and-pencil tests such as those in *Griggs* and *Albemarle.* Supreme Court decisions, however, have in no way limited employee selection devices to paper-and-pencil testing;[104] the Court has consistently suggested a broad definition of testing. In *Griggs,* the Court referred to "testing or measuring procedures,"[105] suggesting that the scope of selection testing extends beyond psychological aptitude or ability tests to other measuring procedures. Moreover, in *Albemarle* the Court expanded the meaning of testing to include employee "tests or selection devices,"[106] again implying that employee selection tests were not limited to paper-and-pencil examinations. In *Albemarle* Justice Stewart also impliedly affirmed the view that the term "testing devices" includes both "testing [and] measuring procedures."[107]

Furthermore, the Supreme Court has shown some deference towards selection tests under the EEOC Guidelines. The Guidelines broadly define tests or selection procedures to include measurements of "physical" requirements.[108] The Guidance Document to the Guidelines

102. *Id.* at 238-39 ("We have never held that the constitutional standard for adjudicating claims of invidious racial discrimination is identical to the standards applicable under Title VII, and we decline to do so today.").

103. *See* Guardians Ass'n v. Civil Serv. Comm'n, 633 F.2d 232, 244-47 (2d Cir. 1980) (*Davis* Court not interpreting "job-relatedness" as defined in Title VII; even if applicable to Title VII determination, *Davis* limited to entry-level tests designed to weed out applicants lacking minimal skills necessary to complete training program), *cert. denied as to Title VII claim,* 463 U.S. 1228 (1983); Craig v. County of Los Angeles, 626 F.2d 659, 663 (9th Cir. 1980) (unlike *Davis,* there was no demonstration of a correlation between test performance and job performance in sheriff's department), *cert. denied,* 450 U.S. 919 (1981); United States v. Virginia, 620 F.2d 1018, 1023 (4th Cir.) (*Davis* confined to fifth amendment principles; Title VII does not require proof of intentional discrimination), *cert. denied,* 449 U.S. 1021 (1980).

104. For example, Justice Stewart speaks of validating "testing *or* measuring procedures" in *Albemarle.* 422 U.S. at 426 (quoting *Griggs,* 401 U.S. at 436) (emphasis added). The conjunction does not seem to be a rhetorical device of repetition, but seems to indicate that types of measurements other than paper-and-pencil tests need to be validated. The EEOC Guidelines and their Guidance Document indicate that validation will apply to measurements such as physical dimensions and biotests. *See infra* notes 108-09 and accompanying text.

105. *Griggs,* 401 U.S. at 436. "[G]ood intent or absence of discriminatory intent does not redeem *employment procedures or testing mechanisms* which operate as 'built-in headwinds' for minority groups and are unrelated to *measuring* job capability." *Id.* at 432 (emphasis added).

106. 422 U.S. at 425.

107. *Id.* at 426 (quoting *Griggs,* 401 U.S. at 436).

108. The term "selection procedure" includes devices that measure "*physical,* educational, and work experience *requirements.*" *See* 29 C.F.R. § 1607.16(Q) (1987) (emphasis added). The Guide-

also supports a broad interpretation of testing, defining employee testing as "selection procedures."[109] Selection procedures include "physical requirements" and "physical job requirements."[110] Biochemical tests, therefore, come within current Supreme Court definitions of selection testing. Because of their inherent inaccuracy, biochemical tests may contribute to unintentional discrimination just as easily as paper-and-pencil aptitude tests.[111]

D. *The Relationship of Tests and Policies.*

In general, a test seeks to measure a specific quality or characteristic. A measurement of a person's vision, height, or weight is a test, of sorts. The connection, however, between the quantitative measurement and the inference drawn from the measurement is so close that one cannot conceive of the two as anything but the same thing. This sort of test is self-validating.

The tests in *Dothard v. Rawlinson*[112]—weight and height—were self-validating. The Court did not require proof of the validity of a weight scale or of a yardstick. These tests were implicitly accepted as valid and accurate. The Court scrutinized the *policy* of requiring a minimum weight and height, not the methods of measurement.

---

lines define "discrimination" as the "use of *any selection procedure* which has an adverse impact on the hiring, promotion, or other employment or membership opportunities of any race, sex, or ethnic group." 29 C.F.R. 1607.3(A) (1987).

109. A document entitled *Adoption of Questions and Answers to Clarify and Provide a Common Interpretation of the Uniform Guidelines on Employee Selection Procedures,* 44 Fed. Reg. 11,995 (1979) [hereinafter Guidance Document] further clarifies the meaning of selection procedures or selection devices:

> 5. Q. Do the Guidelines apply only to written tests?
>
> A. No. They apply to *all selection procedures* used to make employment decisions, including interviews, review of experience or education from application forms, work samples, *physical requirements,* and evaluations of performance. Sections 2B and 16Q, and see Question 6.
>
> 6. Q. What practices are covered by the Guidelines?
>
> A. The Guidelines apply to employee selection procedures which are used in making employment decisions, such as hiring, retention, promotion, transfer, demotion, dismissal or referral. Section 2B. *Employee selection procedures* include job requirements (physical, education, experience), and evaluation of applicants or candidates on the basis of application forms, interviews, performance tests, paper and pencil tests, performance in training programs or probationary periods, and any other procedures used to make an employment decision whether administered by the employer or by an employment agency. *See* Section 2B.

Id. at 11,997.

110. *Id.*

111. The Supreme Court's recent decision in Watson v. Fort Worth Bank & Trust, 108 S. Ct. 2777, 2783-87 (1988), extended disparate impact testing to subjective employment practices. This extension of *Griggs* and *Albemarle* buttresses the above argument for considering biochemical tests as objective selection tests.

112. 433 U.S. 321 (1977); *see also supra* notes 79-85 and accompanying text.

Other tests and the inferences drawn from them are not so closely related. Before a test such as a scholastic aptitude test achieves credibility and efficacy it must be psychologically valid. This validity is often dependent on certain underlying cultural and social assumptions about what is considered intelligence or academic skill. In the employment context, they are typically assumptions about safety, efficiency and profit maximization. These assumptions are why the *Griggs* and *Albemarle* Courts were concerned with the validity of the paper-and-pencil tests. Not only must the quality measured be job-related, but the method of measuring that quality must be accurate. Otherwise, job-relatedness becomes irrelevant.

Most importantly, a test is something quite distinct from the policy it implements. When the two are used in tandem, a court must regard the test and the policy as two discrete elements; the test must be accurate, and the policy it enforces must be job-related. Typically, validation of the test—proof of its accuracy—will also justify the policy as one related to the job. This is especially true when an employer uses the content variety of validation under the Uniform Guidelines. In the case of biochemical tests, however, the potential for inaccuracy compels a two-part analysis: first, an employer must prove the accuracy of the test, and second, the employer must prove the validity of the policy the test implements.

V. VALIDATING BIOCHEMICAL TESTS.

A. *The Advantages of Validation Under the Uniform Guidelines.*

Because of the potential inaccuracies, courts should require employers to validate biochemical tests under a system of demanding and reliable rules to ensure their accuracy and job-relatedness. In *Davis* the Court stated: "[T]here is no single method for appropriately validating employment tests for their relationship to job performance."[113] In spite of the *Davis* aberration, *Griggs* and *Albemarle* are binding in the testing context. The Uniform Guidelines must serve as at least the basis for validating biochemical tests.

The Uniform Guidelines offer several advantages as a method of validation. For example, the Guidelines provide administrative expertise and expert study in the technical and specialized area of testing. The Guidelines are predictable and reliable; courts can point to the same objective standard in each case of biochemical testing. Once an employer has validated a biochemical test under the Guidelines, she can use the

113. 426 U.S. 229, 247 n.13 (1976).

test with confidence. Without a consistent method, the employer would have less confidence in her selection device and might forego drug or AIDS testing altogether. A truly nondiscriminatory test can increase the productivity and safety of the employer's business.

Other advantages are administrability and efficiency. Courts and employers alike can consult the Guidelines rather than groping for some vague notion of safety or efficiency. If courts allow only a safety and efficiency standard and do not focus on the accuracy of the test, the potential for discrimination on the chemical level will always be present. Tests of confirmation must be used in addition to a validation study to eliminate the possibility of chemical discrimination.

Finally, the Guidelines can make validation easier for an employer. The Guidelines allow "transportation" of the validity studies of others.[114] That way an employer would not have to conduct his own expensive study if another employer had already done a similar study. Also, the Guidelines allow an employer to replace the challenged test with a nondiscriminatory alternative procedure.

B. *Validation: A Paradigm.*

An employer should conduct a validation study before implementing a biochemical test as a selection device. If a court finds that the employer has not conducted a validation study, it should compel validation under the Uniform Guidelines.

A court that is assessing the sufficiency of an employer's validation study should begin by determining whether the employer retained the necessary data and records. First, a court must require evidence of the accuracy of the test—proof of the use of confirmation tests, records of cutoff levels and other technical elements of the test. Second, a court should require records of test results.[115] The results should be tabulated for minorities and non-minorities. Third, a court should require the employer to retain data on the safety-sensitivity of the employees' job.[116]

Next the court should determine whether the alternatives of transportation or modification are available. Transportation refers to the borrowing of a validity study conducted by another employer, provided the

114. "Transportability" refers to the use of validity studies not conducted by the user. *See* 29 C.F.R. § 1607(A) (1987). Transporting the validity studies of another employer or an independent scientific group to a similar work setting avoids costs connected with validating. Conceivably, if a test became generally acceptable, there would be virtually no cost in validating by transporting.

115. The test's user should maintain records which will be available for inspection to determine the presence of adverse impact. *Cf.* 29 C.F.R. § 1607.4(A) (1987). The records should indicate such categories as race, sex and ethnicity. *Cf. id.* § 1607.4(B).

116. Safety-sensitive jobs would require criteria of "low error rates" and "high productivity." *See infra* note 124 and accompanying text.

tests and job categories are sufficiently similar.[117] Modification refers to the simple replacement of the discriminatory selection test with a less discriminatory alternative.[118] If transportation and modification are unavailable as alternatives, then a validation study will be necessary to rebut the prima facie case. The validation study should be of the criterion-related variety.[119]

When validation is necessary, the court should require initially that the employer introduce evidence of the accuracy of the test. The method of obtaining the sample should be noted, as well as the cutoff score used

117. *Cf.* 29 C.F.R. § 1607.7 (discussing use of validity studies conducted by another employer or manufacturer).

118. *Cf. id.* § 1607.6(A), (B)(2) (discussing modification or replacement with less discriminatory alternatives).

119. The Guidelines define and describe three important varieties of validation studies: criterion-related validity studies, content validity studies, and construct (trait) validity studies. *Id.* § 1607.14. In addition to these, an employer may use "other professionally acceptable techniques." *Id.* § 1607.14.

Criterion-related studies simply correlate performance on selection tests with any number of criteria that are presumed to measure successful work performance. Examples of criteria include "production rate, error rate, tardiness, absenteeism, and length of service." *Id.* § 1607.14(B)(3). Criterion-related studies may be either predictive, dealing with performance on criteria longitudinally (i.e., over time), or concurrent, dealing with present performance as measured by the criteria. *Id.* § 1607.14(B)(4). A central requirement is differential validity; the pool of candidates for the study must be a representative sample of the candidates normally available in the relevant job market. *Id.* In other words, the candidates studied should represent a cross-section of the races, ethnic groups and sexes in the employable population. Any correlations of criteria to performance also should be statistically significant. *Id.* § 1607.14(B)(5) (setting the significant level of correlation at the 0.05 level of significance, meaning a probability of no more than one in twenty that the correlation occurred by chance). Finally, the validity study should determine whether cutoff scores (for example, the level at which urine tests are considered positive for marijuana) are appropriately set. *Id.* § 1607.14(B)(6).

A content validity study will apply to tests that measure knowledge, skills or abilities important to performance of the job. *See generally id.* § 1607.14(C). The classic content-based test is a typing examination for speed when the job in question is for a typist. Practicable tests for these purposes would be tests of performance.

Construct (trait) validity studies use scores on tests to enforce a policy. An intelligence test can function as a construct-related test. These studies depend on the identification or development of a test that measures or allows the inference of the construct. *See id.* § 1607.14(D)(2) (describing the necessary "job analysis"). Examples of constructs would include traits such as apathy, anti-sociality, tendency to break the law, tendency to use illegal substances, or tendency to fraternize with criminals. *Cf. id.* (job analysis should reflect the "work behavior(s) required for successful performance of the job"). If an employer could show that urine tests for marijuana correlated positively to one of these traits, then the test would validly imply the negative traits under construct validation. In any case, the Guidelines establish an important limitation on the use of construct validity studies: until the professional literature provides more guidance on the use of construct validity studies, such studies will not be acceptable unless preceded by transportable criterion-related validity studies. *Id.* § 1607.14(D)(4)(a). Therefore, an employer would need to have proven that a positive urine test was validated to the construct of criminality or drug use by using studies drawn from the professional literature.

to call a result "positive."[120] A cutoff score set too low may invalidate the job-relatedness of the test by introducing the possibility of detecting cross-reactive compounds or the off-duty use of marijuana in the case of urine testing.[121] The court should ascertain whether the employer used a more specific test to confirm the screening test.[122]

The employer should undertake a job analysis for each affected job category. This analysis should describe duties, safety-sensitivity, work behaviors, and work outcomes of the job.[123] Furthermore, specific criteria for each job category should be named and rated, including but not limited to production rate, error rate, tardiness, absenteeism and length of service.[124] These normative categories should be tabulated for both minorities and non-minorities, so that the requirement for differential validation can be met.[125] Steps should be taken to formalize these criteria on standardized forms in an effort to minimize supervisor (i.e., "rater") bias.[126]

A court might then determine whether employees submitted to a voluntary program of random drug testing. The employees chosen should reflect a random cross-section or representative sample of ethnic and racial categories as well as positions in the line of career progression.[127]

Statisticians should assess the correlation between job performance and marijuana levels in the urine or the numerical value of the AIDS test. If a positive correlation between measurements exists, then the employer may present these findings to the court as evidence that the biochemical test is in compliance with the Guidelines.[128]

Even if a validation study establishes the accuracy of a biochemical test, the test still may not be job-related. This is especially true in the

120. *See supra* notes 76-78 and accompanying text; *see also* 29 C.F.R. § 1607.14(B)(6) (discussing cutoffs).

121. *See supra* notes 29-32 and accompanying text.

122. *See supra* notes 33-38 and accompanying text.

123. This study would be a criterion-related study of the kind described in 29 C.F.R. § 1607.9(A). *See supra* note 118.

124. *Cf.* 29 C.F.R. § 1607.14(B)(3).

125. *Id.* §§ 1607.4(A), (B).

126. *See* Albemarle Paper Co. v. Moody, 422 U.S. 405, 432 (1975) (quoting 29 C.F.R. § 1607.4(B)(3), (4) (1970), which deal with standardizing appraisal forms and scrutinizing supervisorial ratings to prevent bias).

127. *Cf. id.* § 1607.14(B)(4).

128. *Id.* § 1607.14(B)(5) (A criterion is related to performance if the relationship is statistically significant; that is, the probability of the positive correlation occurring by chance is no more than one in twenty.). Ideally, this criterion-related study should be replaced by a construct validity study when the professional literature is sufficiently developed on the subject. *See supra* note 121 and accompanying text (reviewing various constructs appropriate for urine testing under 29 C.F.R. § 1607.14(D)).

case of AIDS testing; the majority of research indicates that, contrary to popular wisdom, one cannot contract AIDS through casual contact.[129] An employee infected with AIDS will perform at a similar physical level until the later stages when brain disease develops. Therefore, even with high cutoff levels and the Western blot assay, AIDS testing may not be job-related except in certain health-related professions and occupations such as the armed forces, where blood donations are required. Validation eliminates the possibility of discrimination and allows the necessary testing in health-sensitive positions.

If an employer proves the accuracy of a urine test for drugs, the employer can probably justify the test on the grounds of safety and efficiency. This rationale is persuasive in safety-sensitive positions, but in non-safety-sensitive positions, a urine test may not be directly related to job performance. The oversensitivity of urine tests makes them especially problematic in non-safety-sensitive positions. An employee could test positive weeks after ingestion of the drug.[130] Arguably, what an employee does during his off-duty hours is a private matter and is not related to job performance. On the other hand, an employee's off-duty conduct may reflect on his business, perhaps damaging goodwill. The inquiry into the job-relatedness of urine testing, is thus fact specific, and can proceed only when the employer has verified the accuracy of the urine test.

VI. CONCLUSION

Employee urine testing should increase disparate impact litigation under Title VII. Socioeconomic patterns of drug use coupled with passive inhalation and the potential for cross-reactivity may cause some ethnic groups to suffer unduly from biochemical drug testing. Given the proper use of statistical data plaintiffs may state a prima facie case and force employers to validate biochemical tests. Cutoff levels set too low will detect passive exposure, thereby differentially excluding some from hiring or promotions.

The proliferation of AIDS testing in the employment setting also should bring an increase in Title VII litigation. In the first instance, blacks and hispanics are afflicted with AIDS at higher rates than whites. Simply because individuals test positive for the AIDS virus does not imply that performance on the job will be affected or that unsafe conditions will result. Only in the late stages of AIDS would emaciation or encephalopathy impair job performance. Likewise, employees with AIDS

129. *See supra* note 78 and accompanying text.

130. *See supra* note 32 and accompanying text.

would not create unsafe conditions for other employees or the public since the virus is not contagious through casual contact. Furthermore, most test results of the ELISA screen for AIDS will be false positive in low-risk populations. If an employer tested protected members of a low-risk population and these applicants were excluded on the basis of unconfirmed ELISA screens, then a claim for disparate impact could follow.

The heightened potential for Title VII litigation underscores the importance of defenses for employers. If biochemical testing is classified as an employee selection policy, then the employer need simply articulate a single nondiscriminatory reason why the policy promotes safety and efficiency. If biochemical testing is classified as an employee selection test, then the employer must either validate, transport validation, or modify the protocols for biochemical testing. The case law, the EEOC Guidelines and the Guidance Document all support the conclusion that biochemical testing should be classified as an employee selection test. Furthermore, analysis of biochemical testing as selection testing offers several policy advantages. First, validation studies will minimize the effects of discriminatory biochemical tests and discriminatory uses of valid biochemical tests. Second, with ready knowledge of a biochemical testing protocol's validity an employer could make employment decisions more confidently. Finally, validation serves the goals of safety, efficiency and profit maximization. A valid selection test can benefit all; the employer is allowed to pursue valid business goals while society protects the rights of all categories of citizens.

Andrew Ayers Martin