



Application of Title VII to Jobs in High Places

Citation

Elizabeth Bartholet, Application of Title VII to Jobs in High Places, 95 Harv. L. Rev. 945 (1982).

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:3138563>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

APPLICATION OF TITLE VII TO JOBS IN HIGH PLACES

Elizabeth Bartholet

TABLE OF CONTENTS

	PAGE
I. INTRODUCTION	948
A. <i>The Problem</i>	948
B. <i>Development and Meaning of the Business Necessity Test on the Lower Level</i>	950
C. <i>Implications for the Upper Level: The Meritocratic Myth</i>	955
II. DEVELOPMENT OF A DIFFERENTIAL STANDARD: FROM JUDICIAL SCRUTINY TO JUDICIAL DEFERENCE	959
A. <i>Judicial Hostility to Title VII Challenges on the Upper Level</i>	959
B. <i>Judicial Distortion of Key Title VII Doctrines in Upper Level Cases</i>	964
1. <i>Experience and Educational Qualifications</i>	965
(a) <i>Limited Burden Imposed on Employers to Show Validity of Qualifications</i>	965
(b) <i>Qualifications Assumed Valid in Analyzing Racial Impact</i> ..	970
2. <i>Subjective Selection Systems</i>	973
3. <i>Tests</i>	978
III. THE CASE FOR A SINGLE STANDARD	978
A. <i>Elitist Nature of the Current Differential Standard</i>	978
B. <i>The 1972 Amendments to Title VII and Their Legislative History</i> ..	980
C. <i>Applicability of Traditional Title VII Standards to Upper Level Selection Systems</i>	983
1. <i>Privacy and Associational Interests</i>	983
2. <i>Difficulties of Demonstrating Job-Relatedness and Business Necessity</i>	984
(a) <i>Nature of the Jobs</i>	984
(b) <i>Nature of the Selection Systems</i>	985
3. <i>Quality of Performance</i>	989
IV. PRACTICAL IMPLICATIONS OF A SINGLE STANDARD: MAJOR ISSUES IN APPLYING TRADITIONAL TITLE VII STANDARDS TO UPPER LEVEL EMPLOYMENT	998
A. <i>Importance of Class Action Treatment</i>	1000
B. <i>Importance of Disparate Impact Analysis as Compared with Disparate Treatment Analysis</i>	1004
C. <i>Advantages of Griggs Approach in Assessing Legality of Subjective Selection Systems</i>	1006
D. <i>Assessing Proof of Job-Relatedness and Business Necessity on the Upper Level</i>	1008
1. <i>Limitations of the Testing Experts</i>	1008
2. <i>Problems and Potential of Validation on the Upper Level</i>	1009
(a) <i>Job Analysis</i>	1009
(i) <i>Inadequacy of Job Description Approach</i>	1010
(ii) <i>Race as Job-Related</i>	1013

(b) <i>Validation Techniques</i>	1016
(i) <i>Content Validity</i>	1016
(ii) <i>Empirical Validity</i>	1018
3. <i>Duty to Explore and Adopt Less Discriminatory Alternatives of</i> <i>Equivalent Validity</i>	1023
V. CONCLUSION	1026

HARVARD LAW REVIEW

APPLICATION OF TITLE VII TO JOBS IN HIGH PLACES

By Elizabeth Bartholet*

Employers of blue collar workers have seen their employment practices exposed to the brilliant light of title VII inquiry. When selection methods involving tests, degree requirements, and the like have effectively shut out minorities, courts have insisted that employers either justify the methods or abandon them. In contrast, when plaintiffs have challenged methods for choosing teachers, lawyers, or managers, courts have left those methods largely in the shadows, unexamined and undisturbed — even when they exclude minorities and are as little justified as the methods condemned at the lower level. Professor Bartholet argues that this emerging distinction between upper and lower level jobs is indefensible as a matter of either law or policy. She concludes that title VII has had a salutary effect on lower level employment and that courts should use title VII to further racial equality at the upper level as well.

THE federal courts have developed standards defining race discrimination in lower level employment that have cast doubt on many of our traditional methods of job allocation. Since the mid-1960's, written tests and informal subjective systems have regularly been ruled unlawful. Indeed, any employment scheme with a racially exclusionary impact has been treated as suspect. As a result, employers have been forced to revise traditional methods of selection, and blacks have made significant gains in employment status.

Many courts appear reluctant to apply comparable standards to the upper level jobs that have increasingly become the focus of litigation. This Article argues that there is no legal basis for distinguishing between upper and lower level selection methods. Application of traditional standards would encourage the full racial integration contemplated by title VII

* Assistant Professor of Law, Harvard University. Radcliffe College, B.A., 1962; Harvard Law School, LL.B., 1965.

I want to thank the many friends and colleagues who have provided generous support and helpful comments on early drafts. I am particularly indebted to Richard S. Barrett, Robert W. Bennett, Richard E. Biddle, Jeanne M. Brett, George Cooper, Harry T. Edwards, Stephen B. Goldberg, Barry L. Goldstein, Duncan M. Kennedy, Lance M. Liebman, Frank I. Michelman, Paul Weiler, and Lloyd L. Wienreb.

of the Civil Rights Act of 1964. It would also expose irrationalities in our job allocation systems and open up possibilities for improved selection.

I. INTRODUCTION

A. *The Problem*

Courts have gone far during the last fifteen years to ensure black participation¹ in the ranks of the work force. Blacks have gained access to industries and crafts traditionally closed to them, and to responsible blue collar supervisory positions. Blacks have made only limited progress, however, in gaining access to "upper level" jobs — jobs with high pay or status, or with significant social or political power.² Relatively few blacks have entered the more elite professions or attained pol-

¹ For simplicity, this Article refers primarily to blacks, but the discussion is intended to apply to other racial or ethnic groups historically subject to extreme intentional discrimination. Although the discussion is limited to racial or ethnic discrimination, it has obvious relevance to other forms of discrimination, such as that based on sex, handicap, or age. Indeed, many of the upper level cases discussed here involved women charging sex discrimination.

It is important to focus on race separately, because the issues are not necessarily the same for all the groups protected by antidiscrimination legislation. For example, the disparate impact doctrine adopted in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1974), is of central significance in upper level employment discrimination. That doctrine was developed in the context of race discrimination and has only recently begun to be applied in other contexts. See, e.g., *Dothard v. Rawlinson*, 433 U.S. 321 (1977) (applying impact doctrine to sex discrimination claim). *Griggs* was predicated on the assumption that blacks and whites are inherently equal in ability and that, but for historical discrimination, they would be equally well situated in employment. See *Griggs*, 401 U.S. at 430-31. Similar assumptions cannot be made for all other protected groups. Indeed, some — such as the aged and the handicapped — may be protected, in part, because they are likely as groups to be less qualified for employment than others. Moreover, one might conclude that the impact doctrine is justified for race because of a combination of factors that include the social importance of achieving a racially integrated work force. The question whether the impact doctrine should be similarly applied in other areas would then turn in part on the perceived importance of achieving full integration of a particular protected group in the work force.

² This Article makes no attempt to define precisely the line between upper and lower level jobs. The argument developed in Part II is that many courts apply differing legal standards, depending on the socioeconomic status of the job involved. Courts have tended to apply strict standards to systems involving blue collar jobs, including supervisory jobs and highly skilled craft jobs. The courts have applied similar standards to systems involving white collar jobs with limited status and power. All such jobs are considered lower level for purposes of this discussion. Courts have tended to apply more relaxed standards to systems involving middle and upper management jobs, professional positions, and other jobs requiring advanced educational degrees. All of these are here termed upper level positions.

icymaking positions at major industrial and financial institutions.

It is not surprising that integration of the work force has taken place primarily on the blue collar level. Aggressive efforts to combat discrimination did not begin until the passage of title VII of the Civil Rights Act of 1964,³ and early efforts were aimed at lower level jobs, which represented the great bulk of employment opportunities. Not until 1972 did Congress expand title VII to cover academic institutions⁴ and public as well as private employment.⁵ Furthermore, blacks have found it harder to satisfy education and experience requirements for upper level jobs than to satisfy similar requirements for less prestigious jobs. Only in the last few years, therefore, have significant numbers of challenges been brought to upper level employment decisions.

Recent legal developments call into question whether the principles applied to open up blue collar jobs to blacks will be applied to upper level jobs.⁶ The courts have tended to show far greater deference to upper than to lower level employers, as is discussed in Part II. While the Supreme Court has not explicitly differentiated between upper and lower level employment systems, it has hinted, just as upper level job issues are coming to the forefront, that it may be prepared to relax title VII standards for employers on all levels.⁷

³ Pub. L. No. 88-352, §§ 701-716, 78 Stat. 241, 253-66 (1964) (codified as amended at 42 U.S.C. §§ 2000e to 2000e-17 (1976 & Supp. III 1979)).

⁴ Equal Employment Opportunity Act of 1972, Pub. L. No. 92-261, § 3, 86 Stat. 103, 103-04 (codified at 42 U.S.C. § 2000e-1 (1976)).

⁵ *Id.* § 11, 86 Stat. at 111 (codified at 42 U.S.C. § 2000e-16 (1976)).

⁶ For discussion of upper level employment discrimination issues, see D. BELL, RACE, RACISM AND AMERICAN LAW § 9.9 (2d ed. 1980); Bardeen, *The Legal Profession: A New Target for Title VII?*, 55 CAL. ST. B.J. 360 (1980); Devine, *Women in the Academy: Sex Discrimination in University Faculty Hiring and Promotion*, 5 J.L. & EDUC. 429 (1976); Friedman, *Congress, the Courts, and Sex-Based Employment Discrimination in Higher Education: A Tale of Two Titles*, 34 VAND. L. REV. 37 (1981); Hunt & Pazuniak, *Special Problems in Litigating Upper Level Employment Discrimination Cases*, 4 DEL. J. CORP. L. 114 (1978); Paone & Reis, *Effective Enforcement of Federal Non-Discrimination Provisions in the Hiring of Lawyers*, 40 S. CAL. L. REV. 615 (1967); Waintroob, *The Developing Law of Equal Employment Opportunity at the White Collar and Professional Level*, 21 WM. & MARY L. REV. 45 (1979); Yurko, *Judicial Recognition of Academic Collective Interests: A New Approach to Faculty Title VII Litigation*, 60 B.U.L. REV. 473 (1980); Note, *Title VII and Employment Discrimination in Upper Level Jobs*, 73 COLUM. L. REV. 1614 (1973); Note, *Tenure and Partnership as Title VII Remedies*, 94 HARV. L. REV. 457 (1980); Note, *Applicability of Federal Antidiscrimination Legislation to the Selection of a Law Partner*, 76 MICH. L. REV. 282 (1977); Comment, *Subjective Employment Criteria and the Future of Title VII in Professional Jobs*, 54 U. DET. J. URB. L. 165 (1976).

⁷ See, e.g., *Texas Dep't of Community Affairs v. Burdine*, 450 U.S. 248 (1981);

Courts have tended to obscure the extent to which they have applied a differential standard, presumably because, as discussed in Section B of Part III, title VII appears to provide no basis for distinguishing between lower and upper level jobs. The message that blacks may be excluded from proportionate participation in upper level jobs has consequently been communicated in a singularly obnoxious form. Because the courts have been reluctant to state that upper level employment should be subject to a more lenient standard, they have tended to endorse racially exclusionary upper level selection systems by finding that these systems meet the stringent standards applied to lower level jobs. Employer policies that exclude blacks from positions of responsibility in the most important institutions in our society have been found reasonable and necessary to maintain the quality of those institutions. Demands for access by blacks have been rejected as equivalent to demands for the lowering of standards.

B. Development and Meaning of the Business Necessity Test on the Lower Level

The issue of how title VII's antidiscrimination principle⁸ should be interpreted in upper level job discrimination cases must be approached with an understanding of how the principle has been interpreted in lower level cases. Only because that principle came to mean more than the mere outlawing of intentional discrimination could any question arise about the propriety of applying essentially the same standards to upper level cases.

Title VII clearly bans intentional discrimination on all job

New York City Transit Auth. v. Beazer, 440 U.S. 568, 582-87 (1979); *Washington v. Davis*, 426 U.S. 229, 248-52 (1976).

It may be significant that some of the Court's most restrictive rulings have been handed down in cases involving upper level job discrimination. *See, e.g.*, *Board of Trustees v. Sweeney*, 439 U.S. 24 (1978), *vacating and remanding per curiam* 569 F.2d 169 (1st Cir. 1978); *National Educ. Ass'n v. South Carolina*, 434 U.S. 1026 (1978), *aff'g mem.* *United States v. South Carolina*, 445 F. Supp. 1094 (D.S.C. 1977). The police department jobs at issue in *Washington v. Davis* did not involve the kind of status typical of most upper level jobs. They did, however, involve considerable power over citizens within the department's jurisdiction, power of which the Supreme Court, located within that jurisdiction, would have been well aware.

It may also be significant that the Court has chosen to uphold affirmative action programs in a case involving blue collar industrial workers, *United Steelworkers v. Weber*, 443 U.S. 193 (1979), while striking down such a program in a case involving medical school applicants, *Regents of the Univ. of Cal. v. Bakke*, 438 U.S. 265 (1978).

⁸ The focus throughout this Article is on title VII, since it is the key federal legislation banning race discrimination in employment.

levels. Radical expansion of the antidiscrimination principle came with judicial development of the disparate impact doctrine. Adopting this doctrine in *Griggs v. Duke Power Co.*,⁹ the Supreme Court ruled that employment policies having a disparate impact on blacks were unlawful unless the employer could show that those policies were job-related and justified by "business necessity." Absence of discriminatory intent on the employer's part was deemed irrelevant.

Employers found it hard to satisfy the *Griggs* standard. An employer could not rest on proof that its selection practices¹⁰ reflected a legitimate business purpose and were consistent with general practice in the relevant industry. Instead, the business necessity doctrine required a demonstration "that the practice [was] necessary to the safe and efficient operation of the business."¹¹ Moreover, job-relatedness had to be demonstrated under strict standards resembling those set by professional associations of industrial psychologists¹² and by the Equal Employment Opportunity Commission (EEOC),¹³ the agency with chief responsibility for enforcing title VII. These standards required at a minimum: (1) a "job analysis" — an examination of the actual tasks involved in the job at issue; (2) a test and scoring scheme designed to predict the ability to perform those tasks; and (3) "validation." Courts have generally held that validation requires either a demonstration that success on the employment test correlates with success on the

⁹ 401 U.S. 424 (1971).

¹⁰ "Selection" is used throughout this Article to cover not only initial hiring, but also other job-allocation and salary-determining decisions covered by title VII, such as promotion and transfer. See generally 42 U.S.C. § 2000e-2 (1976) (comprehensive description of unlawful employment practices).

¹¹ *Robinson v. Lorillard Corp.*, 444 F.2d 791, 798 (4th Cir.), cert. dismissed, 404 U.S. 1006 (1971).

¹² Most influential is the American Psychological Association (APA), and its Division of Industrial-Organizational Psychology (Division 14). See AMERICAN PSYCHOLOGICAL ASS'N, AMERICAN EDUC. RESEARCH ASS'N & NAT'L COUNCIL ON MEASUREMENT IN EDUC., STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTS AND MANUALS (rev. ed. 1974) (1st ed. 1966) [hereinafter cited as APA STANDARDS]; DIVISION OF INDUSTRIAL-ORGANIZATIONAL PSYCHOLOGY, AMERICAN PSYCHOLOGICAL ASS'N, PRINCIPLES FOR THE VALIDATION AND USE OF PERSONNEL SELECTION PROCEDURES (2d ed. 1980) (specifically addressing employment selection issues) [hereinafter cited as DIVISION 14 PRINCIPLES]. For a discussion of professional standards and their development, see B. SCHLEI & P. GROSSMAN, EMPLOYMENT DISCRIMINATION LAW 68-70 (1976).

¹³ The 1970 EEOC Guidelines on Employee Selection Procedures, 35 Fed. Reg. 12,333 (1970) (original version promulgated Aug. 24, 1966), have been superseded by the Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. §§ 1607.1 to 1607.18 (1980) (first published at 43 Fed. Reg. 38,290 (1978)) [hereinafter cited as Uniform Guidelines]; the latter were promulgated in 1978 by the four federal agencies chiefly involved in enforcement of title VII.

job ("empirical validity"), or a demonstration that the test itself is so comparable to the actual job that it constitutes an adequate sample of job performance ("content validity").¹⁴ Finally, even if a selection system was valid, it could be held unlawful if there was a comparably valid system with less adverse racial impact.

Application of these standards resulted in the outlawing of challenged employment tests on a nearly wholesale basis during the late 1960's and the early 1970's. Almost all written tests had an adverse impact on blacks, and relatively few such tests had been adopted in accord with the tenets of the industrial psychologists. Employers, for the most part, had never tried to articulate their job performance goals in a systematic fashion, to develop selection devices carefully targeted to serve those goals, or to measure the success of such devices by validity studies. And employers were not in the habit of examining alternative selection devices to discover which would have the smallest adverse impact on blacks while still ensuring a qualified work force.

The *Griggs* doctrine was applied to all objective selection devices commonly used on the lower level. Many cases focused on written tests, because such tests posed a major barrier to black employment. And the courts went furthest in demanding evidence of validation in connection with written tests. But the courts looked with like suspicion at other objective criteria that had a racial impact, such as employer requirements that employees have prior related work experience, that they be free of arrest or conviction records, or that they have high school diplomas. Such requirements were struck down unless the employer demonstrated that they were required by "business necessity."¹⁵

Courts also looked with suspicion at the subjective selection systems commonly used to hire and promote workers on the blue collar level. These systems were almost uniformly condemned, usually on the theory that they lent themselves to the expression of conscious or unconscious bias.¹⁶

Specific cases illustrate the real-life meaning of the legal standards that courts imposed in lower level job cases. In *Griggs*, the defendant company was involved in the generation and distribution of electric power. At issue were jobs as laboratory technicians and jobs in departments dealing with generating and other electrical equipment, crushers, and other

¹⁴ See generally *infra* pp. 1016-23.

¹⁵ See *infra* pp. 965-67.

¹⁶ See *infra* pp. 973-75.

heavy machinery. The employer required that applicants for promotion or transfer either have a high school degree or equivalency certification, or achieve certain scores on mental and mechanical aptitude tests; the cutoff on the mental aptitude test was at approximately the level expected of an average high school graduate. The employer claimed that these requirements were justified by the growing complexity of the work involved; the company also argued that the requirements were essential for jobs at the top of the various lines of progression and that it was important to morale and efficiency that workers have the capacity to rise to the top.¹⁷ Noting that the general intelligence test was designed to measure the "ability to understand, to think, to use good judgment" and that the mechanical ability test was designed to assess understanding of simple machines, the district court found these to be qualities that the defendant "would logically want to find in his employees."¹⁸

There was ample room for argument in *Griggs* that the company's requirements were motivated not by any legitimate business purpose, but rather by illicit discrimination; however, the Supreme Court specifically refused to rely on such an argument. Assuming employer good faith, as did the Court, an uncritical commonsense approach would seem to have compelled a finding that the requirements were job related. Certainly the Supreme Court did not know enough about the running of a power plant to find the employer's arguments irrational. And it was a common business practice to use entry level tests to select employees suitable for future promotion. But the Supreme Court rejected defenses based on apparent commonsense and rationality, finding that acceptance of such defenses would be incompatible with congressional intent.¹⁹ Instead, the Court demanded evidence demonstrating job-relatedness and business necessity,²⁰ presumably because the social cost of excluding blacks from employment was thought to warrant this higher standard of justification.

A recent Second Circuit case, *Grant v. Bethlehem Steel Corp.*,²¹ illustrates the courts' similarly strict approach to subjective selection procedures that have an adverse racial impact. At issue was the system for selecting foremen to supervise ironworkers engaged in the construction of steel frameworks

¹⁷ 420 F.2d 1225, 1231, 1232 n.2 (4th Cir. 1970), *rev'd*, 401 U.S. 424 (1971).

¹⁸ 292 F. Supp. 243, 250 (M.D.N.C. 1968).

¹⁹ 401 U.S. at 432, 436.

²⁰ *Id.* at 431.

²¹ 635 F.2d 1007 (2d Cir. 1980), *cert. denied*, 101 S. Ct. 3083 (1981).

for bridges, skyscrapers, air terminals, and the like. The district court upheld the system, noting that foremen must necessarily be hired according to subjective evaluations of their ability to promote safety and productive work, because there were no readily identifiable objective criteria for determining this ability.²² The Second Circuit reversed and remanded, holding that the employer nonetheless had to demonstrate the business necessity of its particular subjective evaluation process.²³

The legal standards developed in these and other lower level job cases have put significant pressure on employers to change their hiring and promotion systems. When courts strike down old systems, they typically order employers to develop job-related selection methods. In addition, courts sometimes order particular changes in existing systems and often order implementation of temporary selection systems designed to remedy the effects of the unlawful system. Finally, the potential for litigation has prompted lower level employers to assess the validity of their systems and to develop new systems more likely to prove valid.

Employers have also been under some pressure to avoid the costs of litigation — or of validation — by adopting selection systems that have no adverse racial impact. Absent racial impact, there is no title VII obligation to validate selection procedures. *Griggs* has thus encouraged employers to develop hiring and promotion systems that select on a racially proportionate basis from among qualified candidates.²⁴ The Supreme Court's approval in *United Steelworkers v. Weber*²⁵ of a race-conscious employment scheme was the logical consequence of the *Griggs* doctrine. Employers — compelled by *Griggs* to

²² See *id.* at 1012.

²³ *Id.* at 1018–19.

²⁴ A number of commentators have argued that the difficulty of satisfying the *Griggs* demand for justification produces pressure for race-conscious hiring. E.g., B. SCHLEI & P. GROSSMAN, *supra* note 12, at 181; Belton, *Discrimination and Affirmative Action: An Analysis of Competing Theories of Equality and Weber*, 59 N.C.L. REV. 531, 551–52 (1980); Fallon, *To Each According to His Ability, from None According to His Race: The Concept of Merit in the Law of Antidiscrimination*, 60 B.U.L. REV. 815, 841, 853 (1980); Freeman, *Legitimizing Racial Discrimination Through Antidiscrimination Law: A Critical Review of Supreme Court Doctrine*, 62 MINN. L. REV. 1049, 1099 (1978); Meltzer, *The Weber Case: The Judicial Abrogation of the Antidiscrimination Standard in Employment*, 47 U. CHI. L. REV. 423, 426, 434 (1980); Schatzki, *United Steelworkers of America v. Weber: An Exercise in Understandable Indecision*, 56 WASH. L. REV. 51, 61–63 (1980); see Comment, *The Business Necessity Defense to Disparate-Impact Liability Under Title VII*, 46 U. CHI. L. REV. 911, 913 (1979).

²⁵ 443 U.S. 193 (1979).

eschew policies that had an unnecessary adverse impact on blacks — could not be penalized for adopting policies designed to ensure that blacks were employed on a proportionate basis.

C. Implications for the Upper Level: The Meritocratic Myth

The radical implications of *Griggs* for upper level employment systems are obvious when one examines the objective criteria and subjective processes that such systems involve. They place a premium on educational accomplishment and prior relevant experience. For candidates who satisfy minimum objective qualifications, the final decisionmaking tends to be largely subjective, based on evaluations of the candidates' previous work and potential for future performance.

Selection systems of this sort are likely to have an adverse racial impact. Blacks as a group are far less likely than whites to have had the education and experience that have traditionally been the prerequisites for these jobs. Use of such credentials — either as minimum objective requirements or as part of a subjective evaluation process — seriously limits black access to upper level positions. Subjective systems may, of course, be used to further “affirmative action” goals. But they also allow for the expression both of conscious bias and of the unconscious bias that is likely to result in the exclusion of persons who are visibly different from those doing the selecting.

Although employers may be able to generate commonsense justifications for their systems and demonstrate that they mirror those of comparable employers, such defenses are insufficient under *Griggs*. Upper level employers will find it hard to provide the proof of job-relatedness and business necessity demanded in lower level cases. Upper level selection devices have only rarely been developed and validated in any kind of systematic manner. Testing experts²⁶ have found them suspect, just as they found suspect the standardized written tests commonly used in lower level selection.²⁷ Thus, if disparate

²⁶ “Testing experts” is used here to indicate the type of people accepted by courts as expert witnesses on employment selection issues. Their training would typically be in industrial psychology, testing and measurement, behavioral science, and related fields.

²⁷ The testing literature indicates that selection on the higher levels is generally done in an extremely unscientific fashion, that research done by industrial psychologists is rarely applied to refine the selection devices actually used, and that the devices most relied on are supported by extraordinarily little evidence of validity. See J. CAMPBELL, M. DUNNETTE, E. LAWLER & K. WEICK, *MANAGERIAL BEHAVIOR, PERFORMANCE, AND EFFECTIVENESS* 127-45, 469-70 (1970) [hereinafter cited as *MANAGERIAL BEHAVIOR*]; see also R. GUION, *PERSONNEL TESTING* 458 (1967) (the

impact on the upper level is to trigger the traditional *Griggs* burden of justification, a significant rethinking of traditional selection methods will be required.

Two questions are central in considering what legal standards should be applied to upper level job selection. The first is whether there are differences between upper and lower level selection that call for differences in legal treatment. Various courts and commentators have argued for relative immunity for upper level employers on the grounds that at this level employees are not fungible, subjective evaluation is essential, and job-relatedness is difficult if not impossible to demonstrate. Part III of this Article argues that differences in legal treatment of upper and lower level selection systems cannot be justified on the basis of such distinctions.

A second question is whether the *Griggs* doctrine has proven inconsistent with meritocratic principles. The doctrine was initially expounded and defended on meritocratic grounds. The avowed goal was to get employers to eliminate the irrational factor of race from employment selection schemes — to do no more than follow their economic self-interest by developing truly merit-based selection systems.²⁸ But if it was not obvious when *Griggs* was decided in 1971, it has become apparent since then that developing “valid” selection systems is no simple matter.²⁹ Some commentators have criticized the *Griggs* approach, arguing that it has resulted in wasteful validation costs,³⁰ a movement away from merit-based employ-

same validation principles apply to managerial selection systems as to other selection systems, although effective performance may be harder to identify and measure on the managerial level); *infra* pp. 986, 988–89.

²⁸ See, e.g., Cooper & Sobol, *Seniority and Testing Under Fair Employment Laws: A General Approach to Objective Criteria of Hiring and Promotion*, 82 HARV. L. REV. 1598, 1678 (1969). This is the leading article setting forth the theoretical groundwork for *Griggs*.

²⁹ Many have commented on the difficulties of satisfying the legal standards established by *Griggs* and related cases. See, e.g., Holt, *A View from Albemarle*, 30 PERSONNEL PSYCHOLOGY 65, 73–77 (1977); Johnson, *Albemarle Paper Co. v. Moody: The Aftermath of Griggs and the Death of Employment Testing*, 27 HASTINGS L.J. 1239 (1976); Wilson, *A Second Look at Griggs v. Duke Power Company: Ruminations on Job Testing, Discrimination, and the Role of the Federal Courts*, 58 VA. L. REV. 844, 873–74 (1972); Note, *Employment Testing and Proof of Job-Relatedness: A Tale of Unreasonable Constraints*, 52 NOTRE DAME LAW. 95, 95–96 (1976); Note, *Business Necessity Under Title VII of the Civil Rights Act of 1964: A No-Alternative Approach*, 84 YALE L.J. 98, 117–18 (1974).

³⁰ On the costs of validation, see Gwartney, Asher, Haworth & Haworth, *Statistics, the Law and Title VII: An Economist's View*, 54 NOTRE DAME LAW. 633, 643 (1979); Lerner, *Employment Discrimination: Adverse Impact, Validity, and Equality*, 1979 SUP. CT. REV. 17, 18 & n.6; White & Francis, *Title VII and the Masters of Reality: Eliminating Credentialism in the American Labor Market*, 64 GEO L.J. 1213, 1234–35 (1976).

ment toward quota hiring, and a consequent lowering of the quality of job performance.³¹

These concerns have led some to call for a general retreat from *Griggs* — a softening of the burden of justification imposed on employers at all job levels. Alternatively, such concerns may be thought to justify development of a two-tier standard. Courts may have felt fairly comfortable with the more radical implications of *Griggs* for lower level jobs because these jobs did not seem particularly important. On the upper level, however, where courts feel that the quality of performance really matters, they may be reluctant to interfere with traditional selection methods.³²

A major thesis of this Article is that the *Griggs* doctrine has not proven inconsistent with meritocratic principles. The mass of employment discrimination litigation over the past fifteen years has done more than merely educate us about the difficulties of validation. It has also exposed to an unprecedented extent the way in which jobs are allocated to various groups in our society. The argument developed in Section C of Part III is that this litigation has helped destroy the notion that the meritocratic principle is the norm governing job allocation. The systems exposed have not been outlawed because employers could not meet an impossibly strict standard of validation; they have been outlawed because they were revealed to be inconsistent with merit selection. Courts' willingness to apply strict standards of validation grew in part from their recognition of how hard it was to justify in meritocratic terms the job allocation systems they observed.

The *Griggs* doctrine, with its demand for proof of business necessity, can and should serve a similar function with respect to upper level employment systems. A standard that looks instead to whether there are any plausible commonsense justifications for employer practices — whether they serve legitimate business purposes — will result in wholesale immunity from scrutiny on this level, as it would have at the lower level. The fact is that there are seeming commonsense justifications for most of our job selection systems. *Griggs* requires employers to make a persuasive case that these justifications are so related to merit and quality of performance that they warrant

³¹ See, e.g., B. SCHLEI & P. GROSSMAN, *supra* note 12, at 181; Fallon, *supra* note 24, at 841, 853; Lerner, *supra* note 30, at 23, 49; Meltzer, *supra* note 24, at 426-27; Comment, *supra* note 24, at 913.

³² Those favoring the application of relaxed legal standards to upper level employment schemes often rely on such rationales. See, e.g., Fallon, *supra* note 24, at 854-60; Olmsted, *Law as a Business: The Impact of Title VII on the Legal "Industry,"* 10 VAL. U.L. REV. 479, 489-90 (1976).

a racially exclusionary effect. *Griggs* offers the only standard that will result in meaningful reexamination of traditional selection systems.

Enforcement of *Griggs* at the upper level would generate pressure for racially proportionate hiring from among the pool of those with conceded minimum qualifications. The argument developed in Parts IV and V is that this is as it should be. Validation and other proof of business necessity are particularly difficult on the upper level, largely because we are unsure what we mean by effective performance in our most important jobs. Our notions of effective performance are necessarily value laden, reflecting what are often essentially political choices. These considerations militate in favor of opening such jobs to groups traditionally excluded from them. Moreover, the "quota" alternative to validation allows employers significant freedom to choose among selection systems. As long as the systems have no adverse racial impact, there is no title VII violation.

Other kinds of arguments for racially proportionate hiring rely on remedial concerns related to the historic injustice of slavery and segregation. Such arguments have been the primary subject of the continuing debate about "affirmative action," which tends to ask the questions whether and how white society should make up to blacks for what it has done to them. Affirmative action programs are generally assumed to involve preferential treatment of those with lesser qualifications, and the question debated is whether such treatment can be justified by remedial goals. The *Griggs* impact doctrine grew in part out of similar concerns: the Supreme Court that struck down tests and high school degree requirements in *Griggs* was influenced by the historic injustices that had prevented blacks from receiving the kind of education that would have enabled them to satisfy such requirements.

The central rationale of the Court's decision in *Griggs*, however, was something quite different. It was based on an assumption that those of different races are inherently equal in ability and intelligence, and on a deep skepticism about the utility of devices traditionally used to select among applicants for employment.³³ This rationale provides an independent ba-

33

[O]n the record in the present case, "whites register far better on the Company's alternative requirements" than Negroes. . . . This consequence would appear to be directly traceable to race. Basic intelligence must have the means of articulation to manifest itself fairly in a testing process. Because they are Negroes, petitioners have long received inferior education in segregated schools What is required by Congress is the removal of artificial, arbitrary, and

sis for many of the programs debated today in the name of affirmative action — a basis significantly less condescending to the programs' purported beneficiaries than notions of "preferential treatment."³⁴ The *Griggs* rationale is also useful in pushing toward the reform of job allocation schemes in ways that have nothing to do with race.

II. DEVELOPMENT OF A DIFFERENTIAL STANDARD: FROM JUDICIAL SCRUTINY TO JUDICIAL DEFERENCE

The goal of this Part is to show that the courts treat upper and lower level jobs differently under title VII. Why a differential standard is being applied, and whether it can be justified, are discussed in Part III.

Because courts generally purport to apply the same standards to all employment systems, it is difficult to document the existence of a differential standard. No claim is made here that all courts are applying such a standard. Indeed, many appear to be trying to apply the traditional standards to cases at both levels. The claim here is simply that the courts have generally granted upper level employers significant immunity from title VII scrutiny.

A. *Judicial Hostility to Title VII Challenges on the Upper Level*

Many judicial opinions convey a strong feeling that upper level systems should be largely immune from title VII coverage. This attitude is immensely important, for courts can implement a differential standard in myriad ways. A court's exercise of discretion in factfinding shapes the legal conclusion that it actually reaches; and in most title VII cases, facts are

unnecessary barriers to employment when the barriers operate invidiously to discriminate on the basis of racial or other impermissible classification.

Griggs, 401 U.S. at 430-31 (quoting opinion below, 420 F.2d 1225, 1239 n.6 (4th Cir. 1970)) (footnote omitted).

³⁴ Professor Derrick Bell has criticized the classic "affirmative action" approach in the context of minority admissions to advanced educational institutions. Bell, *Bakke, Minority Admissions, and the Usual Price of Racial Remedies*, 67 CALIF. L. REV. 3 (1979). He notes that the general solution has been to use minority racial status as a positive factor in existing admissions systems rather than to reformulate admissions standards — a solution that "has served to validate and reinforce traditional policies while enveloping minority applicants in a cloud of suspected incompetency":

Even the term "affirmative action," which encompasses minority admissions programs, connotes the undertaking of remedial activity beyond what normally would be required. It sounds in *noblesse oblige*, not legal duty, and suggests the giving of charity rather than the granting of relief.

Id. at 8.

determined by judges rather than juries.³⁵ Of similar importance is court discretion in such procedural matters as class action certification, preliminary relief, and discovery. Courts hostile to the merits of title VII claims will tend, whether intentionally or not, to decide key factual and procedural issues in ways that provide immunity to employers.

Judicial hostility to title VII's application on the upper level is apparent in a recent decision holding a law firm's partnership selection process entirely immune from legislative prohibitions of discrimination.³⁶ In a case involving King and Spalding, an Atlanta law firm with over one hundred attorneys, the court found dispositive the existing partners' interests in privacy and freedom of association:

In a very real sense a professional partnership is like a marriage. It is, in fact, nothing less than a "business marriage" for better or worse. Just as in marriage different brides bring different qualities into the union — some beauty, some money, and some character — so also in professional partnerships new mates or partners are sought and betrothed for different reasons and to serve different needs of the partnership. . . . In both, new mates are expected to bring not only ability and industry, but also moral character, fidelity, trustworthiness, loyalty, personality and love. Unfortunately, however, in partnerships, as in matrimony, these needed, worthy and desirable qualities are not necessarily divided evenly among the applicants according to race, age, sex or religion, and in some they just are not present at all. To use or apply Title VII to coerce a mismatched or unwanted partnership too closely resembles a statute for the enforcement of shotgun weddings.³⁷

A number of courts have suggested that title VII, although still applicable, should be relaxed for upper level employers.³⁸

³⁵ Title VII has been interpreted not to provide a right to jury trial even when back pay claims are involved. See, e.g., *Slack v. Havens*, 522 F.2d 1091 (9th Cir. 1975); *EEOC v. Detroit Edison Co.*, 515 F.2d 301 (6th Cir. 1975). *vacated on other grounds and remanded*, 431 U.S. 951 (1977).

³⁶ *Hishon v. King & Spalding*, 24 Fair Empl. Prac. Cas. (BNA) 1303 (N.D. Ga. 1980). *Contra Lucido v. Cravath, Swaine & Moore*, 425 F. Supp. 123, 126 (S.D.N.Y. 1977) ("Congress clearly included in the objectives of Title VII the elimination of job discrimination in professional fields including law and medicine.").

³⁷ *Hishon v. King & Spalding*, 24 Fair Empl. Prac. Cas. (BNA) at 1304-05.

³⁸ See, e.g., *Townsend v. Nassau County Medical Center*, 558 F.2d 117, 120 (2d Cir. 1977) (standards for validating advanced degree requirements for professional jobs should be relaxed), *cert. denied*, 434 U.S. 1015 (1978); *Vuyanich v. Republic Nat'l Bank*, 505 F. Supp. 224, 370-75 (N.D. Tex. 1980) (validation standards should be relaxed for upper level bank jobs).

The *Townsend* court relied on *Spurlock v. United Airlines, Inc.*, 475 F.2d 216,

And some courts have overtly applied a differential standard. In cases involving academic institutions, a number of courts have adopted a "hands-off" doctrine,³⁹ which the Second Circuit has characterized as an "anti-interventionist policy [that] has rendered colleges and universities virtually immune to charges of employment bias, at least when the bias is not expressed overtly."⁴⁰ The hands-off doctrine has been criticized as a form of judicial abdication, inconsistent with the 1972 amendments to title VII that specifically removed the exemption for academic institutions.⁴¹ But the doctrine persists.⁴² Courts often seem outraged at being forced to apply title VII to academic cases; they argue implicitly or explicitly for a more relaxed standard. For example, the Second Circuit recently complained bitterly about the length of tenure case trials: "We do not understand how either the federal courts or the universities can operate if [adverse tenure decisions] are regularly taken to court and entail burdens such as those here incurred."⁴³ In another academic case, a district court concluded that applying traditional title VII standards would be practically equivalent to taking over promotion and tenure decisions.⁴⁴

219 (10th Cir. 1972), an oft-cited case involving airline pilot selection, in which the 10th Circuit proposed a sliding-scale standard under which validation requirements would be relaxed in relation to the importance of the job.

³⁹ In *Faro v. New York Univ.*, 502 F.2d 1229 (2d Cir. 1974), a sex discrimination case, the court stated: "Of all the fields, which the federal courts should hesitate to invade and take over, education and faculty appointments at a University level are probably the least suited for federal court supervision." *Id.* at 1231-32. The opinion scorned the plaintiff's claim: "Dr. Faro, in effect, envisions herself as a modern Jeanne d'Arc fighting for the rights of embattled womanhood on an academic battlefield, facing a solid phalanx of men and male faculty prejudice." *Id.* at 1231. In *EEOC v. Tufts Inst. of Learning*, 421 F. Supp. 152 (D. Mass. 1975), the court laid out a standard completely at odds with lower level cases: it stated that plaintiffs could prevail only if they showed intentional bias or "no rational basis" for a selection policy. *Id.* at 158. Comparable deference was afforded academic employers in *Peters v. Middlebury College*, 409 F. Supp. 857 (D. Vt. 1976); *Labat v. Board of Higher Educ.*, 401 F. Supp. 753 (S.D.N.Y. 1975); *Pace College v. Commission on Human Rights*, 38 N.Y.2d 28, 339 N.E.2d 880, 377 N.Y.S.2d 471 (1975).

⁴⁰ *Powell v. Syracuse Univ.*, 580 F.2d 1150, 1153 (2d Cir.), *cert. denied*, 439 U.S. 984 (1978).

⁴¹ *See, e.g., Davis v. Weidner*, 596 F.2d 726, 731-32 (7th Cir. 1979); *Powell v. Syracuse Univ.*, 580 F.2d 1150, 1153 (2d Cir.), *cert. denied*, 439 U.S. 984 (1978); *Sweeney v. Board of Trustees*, 569 F.2d 169, 176 (1st Cir.) (noting "misgivings over one theme recurrent in these opinions: the notion that courts should keep 'hands off' the salary, promotion, and hiring decisions of colleges and universities"), *vacated and remanded per curiam*, 439 U.S. 24 (1978).

⁴² *See generally Yurko, supra* note 6, at 474, 482-83, 490-91, 495-99 (courts apply far more deferential standard to academic institutions than the law warrants).

⁴³ *Lieberman v. Gant*, 630 F.2d 60, 62 n.1 (2d Cir. 1980).

⁴⁴ *Johnson v. University of Pittsburgh*, 435 F. Supp. 1328, 1371 (W.D. Pa. 1977).

In most upper level cases, traditional title VII standards are purportedly applied, but liability is generally denied. A review of large numbers of these cases leaves a strong impression that courts simply look at upper level cases with different eyes.

Courts in lower level cases have traditionally been willing to assess candidates' qualifications in order to resolve discrimination claims, even when these assessments seemed difficult. In *Grant v. Bethlehem Steel Corp.*,⁴⁵ the Second Circuit analyzed the qualifications of blacks and Puerto Ricans seeking jobs as foremen supervising steel construction workers. The court decided that the minority candidates were qualified, despite the employer's claim that its subjective evaluations were essential to select those capable of adequate performance.

In upper level cases, by contrast, the courts often profess their lack of expertise and refuse to assess candidates' qualifications.⁴⁶ The courts also deny litigants the discovery that would make such assessments possible.⁴⁷ As a result, it is virtually impossible to win any discrimination case brought on an individual basis, because proof in such cases has traditionally involved a showing that the employer has subjected comparably qualified individuals of different races to disparate treatment.

Judicial reluctance to intervene in upper level employment decisionmaking is also apparent in the way courts deal with a variety of key procedural issues. In considering plaintiffs' rights to relief, courts often seem to weigh the equities quite differently on the upper than on the lower level.⁴⁸ In lower

⁴⁵ 635 F.2d 1007 (2d Cir. 1980), *cert. denied*, 101 S. Ct. 3083 (1981).

⁴⁶ See, e.g., *Johnson v. University of Pittsburgh*, 435 F. Supp. 1328, 1371 (W.D. Pa. 1977); *Cussler v. University of Md.*, 430 F. Supp. 602, 605-06 (D. Md. 1977).

⁴⁷ In academic cases, courts have refused to allow plaintiffs access to information regarding the qualifications of other candidates that might be essential to proving liability. See, e.g., *Lieberman v. Gant*, 630 F.2d 60 (2d Cir. 1980); *Keyes v. Lenoir Rhyne College*, 552 F.2d 579, 581 (4th Cir.), *cert. denied*, 434 U.S. 904 (1977); *McKillop v. Regents of the Univ. of Cal.*, 386 F. Supp. 1270, 1278 n.13 (N.D. Cal. 1975).

In a recent case concerning a failure to grant academic reappointment, the court held that plaintiff had no right to discover the votes of two of the named defendants on the reappointment issue. *Gray v. Board of Higher Educ.*, No. 79 Civ. 0062 (S.D.N.Y. Nov. 9, 1981) (denying motion under FED. R. CIV. P. 37). *Contra In re Dinnan* (*Blauberger v. Regents of Univ. Sys.*), 661 F.2d 426 (5th Cir. 1981). Discovery of how negative employment decisions are made is a normal prerequisite to challenging such decisions.

⁴⁸ See Note, *Employment Discrimination Suits by Professionals: Should the Reinstatement Remedy Be Granted?*, 39 U. PITT. L. REV. 103 (1977) (comparing court reluctance to order reinstatement in professional cases to court provision for reinstatement as matter of course in lower level cases).

level cases, once liability is determined, the victims of discrimination have had a right to "make-whole relief" that has included placement in the job positions they would have had but for discrimination.⁴⁹ Nonetheless, in a case involving a discriminatory discharge of a senior advertising executive, the district court found liability but denied reinstatement. The court relied on arguments that the job at issue involved personal contact with clients and close relationships with top executives.⁵⁰

Further, courts have been restrictive in class certification decisions on the upper level,⁵¹ in sharp contrast to the liberal approach to class certification in lower level cases. There courts have recognized that class treatment is often essential to the vindication of substantive rights.⁵² As set forth in detail in Section A of Part IV, class action treatment is even more important in dealing with upper level job discrimination.

Courts occasionally deny class action treatment in upper level cases on the ground that the proposed class is too small.⁵³ But the central question posed by the numerosity requirement of the federal class action rule is the *practicability* of joinder as an alternative to class treatment.⁵⁴ In lower level employment cases, courts have recognized that joinder of even a small number of plaintiffs may be impracticable when members of the class might have reason to fear coming forward to join an

⁴⁹ The principle of make-whole relief was affirmed in *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 415-22 (1975).

⁵⁰ *EEOC v. Kallir, Philips, Ross, Inc.*, 420 F. Supp. 919, 923-24, 926-27 (S.D.N.Y. 1976), *aff'd mem.*, 559 F.2d 1203 (2d Cir.), *cert. denied*, 434 U.S. 920 (1977).

⁵¹ *See, e.g., Scott v. University of Del.*, 601 F.2d 76, 85-89 (3d Cir.), *cert. denied*, 444 U.S. 931 (1979); *Tuft v. McDonnell Douglas Corp.*, 581 F.2d 1304 (8th Cir. 1978); *Peterson v. Albert M. Bender Co.*, 75 F.R.D. 661 (N.D. Cal. 1977); *Kinsey v. Legg, Mason & Co.*, 60 F.R.D. 91 (D.D.C. 1973) (denying class action certification), *rev'd on other grounds sub. nom. Kinsey v. First Regional Sec. Co.*, 557 F.2d 830, 839 (D.C. Cir. 1977).

⁵² *See, e.g., Bridgesmith, Representing the Title VII Class Action: A Question of Degree*, 26 WAYNE L. REV. 1413 (1980); *Rutherglen, Title VII Class Actions*, 47 U. CHI. L. REV. 688, 706-13 (1980).

The Supreme Court's decision in *East Tex. Motor Freight Sys. v. Rodriguez*, 431 U.S. 395 (1977), has caused some courts to moderate their willingness to grant liberal class action treatment. But *Rodriguez*' significance is unclear, and it seems to represent at most a rejection of an extreme presumption in favor of class certification. *See Bridgesmith, supra*, at 1431-32; *Rutherglen, supra*, at 723.

⁵³ *See, e.g., Tuft v. McDonnell Douglas Corp.*, 581 F.2d 1304, 1308 (8th Cir. 1978); *Scott v. University of Del.*, 601 F.2d 76, 88-89 (3d Cir.), *cert. denied*, 444 U.S. 931 (1979); *Peterson v. Albert M. Bender Co.*, 75 F.R.D. 661, 667 (N.D. Cal. 1977); *Kinsey v. Legg, Mason & Co.*, 60 F.R.D. 91, 100 (D.D.C. 1973), *reversed on other grounds sub nom. Kinsey v. First Regional Sec. Co.*, 557 F.2d 830 (D.C. Cir. 1977).

⁵⁴ *See FED. R. CIV. P. 34(a)(1)*, 95 Harv. L. Rev. 963 1981-1982

action.⁵⁵ A comparable standard would justify class treatment in many upper level cases. Job candidates at the managerial and professional level are under strong pressures to avoid being labeled troublemakers. An individual decision to engage actively in litigation will usually be highly risky, given the low probability of prevailing and the danger of alienating present and potential future employers. Joinder cannot, therefore, be assumed "practicable" solely because the number of potential class members is small.

Courts denying class certification have also relied on the subjective nature of challenged upper level employment practices. They have accepted the arguments of defendants that, because each employment decision is made on the basis of an individual assessment, each decision is different and class treatment is therefore inappropriate.⁵⁶ Such decisions are inconsistent with judicial treatment of class action challenges to lower level subjective systems. The courts have as a matter of course analyzed the legality of such systems *as systems* and have granted class action certification as a means to that end.⁵⁷

B. Judicial Distortion of Key Title VII Doctrines in Upper Level Cases

Courts in upper level cases have systematically distorted or failed to apply two doctrines that have been central in title VII enforcement. The "disparate treatment" doctrine holds that differential treatment of comparably qualified persons of different races justifies an inference of intentional discrimination, unless the employer produces an acceptable explanation. If the employer produces such an explanation, plaintiffs can prevail only if they demonstrate that the employer's explanation is a pretext shielding a discriminatory motive. The "disparate impact" doctrine avoids issues of discriminatory motive; it holds that the use of employment tests and other selection devices that have an adverse racial impact constitutes unlawful discrimination, regardless of the employer's good faith, unless the employer can demonstrate the job-relatedness and business necessity of such devices.

⁵⁵ See, e.g., *Brown v. ARA Servs.*, 13 Fair Empl. Prac. Cas. (BNA) 1721, 1726 (D. Md. 1974).

⁵⁶ See, e.g., *O'Connell v. Teachers College*, 63 F.R.D. 638 (S.D.N.Y. 1974). *Contra Presseisen v. Swarthmore College*, 71 F.R.D. 34, 44 n.11 (E.D. Pa. 1976) (*O'Connell* "analysis improperly focuses on the defendant's defense rather than class claims" (emphasis in original)), *aff'd mem.*, 582 F.2d 1275 (3d Cir. 1978).

⁵⁷ See, e.g., *Johnson v. Georgia Highway Express, Inc.*, 417 F.2d 1122, 1124 (5th Cir. 1969); *Bowe v. Colgate-Palmolive Co.*, 416 F.2d 711, 719-21 (7th Cir. 1969).

The two doctrines are closely related. When an employment *system* is challenged, each doctrine ordinarily involves an initial showing by plaintiffs that the system has an adverse racial effect. And the same showing of adverse effect that constitutes a *prima facie* disparate treatment case may also trigger the *Griggs* disparate impact test.⁵⁸

1. *Experience and Educational Qualifications.* — How courts view the use of experience and educational qualifications is critical to the significance of title VII at the upper level. As noted above, such qualifications play a large role in upper level employment selection both as minimum eligibility requirements and as factors considered in overall subjective assessments.

(a) *Limited Burden Imposed on Employers to Show Validity of Qualifications.* — Lower level employers are subject to a heavy burden of justification for all objective criteria that have a disparate impact, including educational and experience qualifications. Courts require direct evidence of job-relatedness and reject defenses based on prevailing industry practice and uncritical commonsense. Thus, lower courts have struck down minimum height and weight requirements for police and fire officers,⁵⁹ a rule requiring the discharge of employees sub-

⁵⁸ *International Bhd. of Teamsters v. United States*, 431 U.S. 324, 335 n.15, 339-40 (1977). The plaintiff's evidence in a disparate treatment case will ordinarily include more than the "effect" evidence indicated in the text. In order to prove discriminatory intent on a systemwide basis, plaintiffs will include, where possible, evidence of individual instances of discriminatory conduct. Moreover, to justify an inference of discriminatory motive, courts may require evidence of more extreme racial impact in a disparate treatment case than would be required to trigger the *Griggs* burden of justification.

Courts and litigators have often failed to understand the close relationship between these two doctrines. For a fuller discussion of this problem, see *infra* pp. 1004-06. This confusion may stem from the fact that the two doctrines developed separately, although simultaneously, to deal with different kinds of problems. The disparate treatment doctrine has generally been used in cases of discrimination against an individual rather than a class. The *Griggs* doctrine, by contrast, has been used primarily in class action challenges to employment policies. The evidence used to make out a *prima facie* case in an individual action brought on a disparate treatment theory would usually not be sufficient to make out a *prima facie Griggs* impact case.

Refusal to analyze a case in terms of disparate impact theory can, of course, operate as a means of dismissing claims if a court does not want to face up to the question whether the employment practices at issue can be justified as job-related. In *Scott v. University of Del.*, 601 F.2d 76 (3d Cir.), *cert. denied*, 444 U.S. 931 (1979), the Third Circuit analyzed the plaintiffs' claim that the university's promotion system discriminated against blacks solely in terms of disparate treatment theory and dismissed the suit, *id.* at 79, even though the district court decision had discussed the plaintiffs' disparate impact claims as well. 455 F. Supp. 1102, 1105 (D. Del. 1978).

⁵⁹ *Davis v. County of Los Angeles*, 566 F.2d 1334, 1342 (9th Cir. 1977), *vacated as moot*, 440 U.S. 625 (1979); *League of United Latin Am. Citizens v. City of Santa*

ject to repeated wage garnishments,⁶⁰ a requirement of related work experience,⁶¹ and high school diploma requirements.⁶² Courts have also demanded a showing of business necessity, i.e., a showing that the criterion at issue is not only job-related, but also essential to safety and efficiency.⁶³

The Supreme Court has been equally exacting. In *Griggs* the Court struck down a high school degree requirement. In *Dothard v. Rawlinson*,⁶⁴ the Court struck down minimum height and weight requirements for prison guards, noting the lack of evidence correlating the requirements with the physical strength necessary for adequate performance.

A leading Fifth Circuit case, *Watkins v. Scott Paper Co.*,⁶⁵ illustrates how strictly the *Griggs* standard has been applied to high school diploma and equivalency requirements for lower level jobs, even those involving significant skill and responsibility. The employer's high school diploma requirement applied to jobs in two categories. The first were "Chemical Recovery" jobs monitoring complex control mechanisms to prevent "explosive situations" from developing in the boilers. The court found that "[f]ailure of the workers to do their job properly would put all of the employees in the entire plant in extreme danger."⁶⁶ The second were "Technical Control" jobs preparing chemicals for the entire laboratory's operations. The lower court found the high school diploma requirement justi-

Ana, 410 F. Supp. 873, 896 (C.D. Cal. 1976); *Officers for Justice v. Civil Serv. Comm'n*, 395 F. Supp. 378, 380-81 (N.D. Cal. 1975). Such criteria have been struck down because of their disparate impact on women, Latinos, and Asians.

⁶⁰ *Johnson v. Pike Corp. of Am.*, 332 F. Supp. 490 (C.D. Cal. 1971).

⁶¹ See *infra* p. 967.

⁶² See, e.g., *James v. Stockham Valves & Fittings Co.*, 559 F.2d 310, 355 (5th Cir. 1977), *rev'g* 394 F. Supp. 434 (N.D. Ala. 1975), *cert. denied*, 434 U.S. 1034 (1978); *Dozier v. Chupka*, 395 F. Supp. 836, 850 (S.D. Ohio 1975). See generally *Hunt & Pazuniak*, *supra* note 6, at 134-35 (discussing courts' tendency to invalidate the high school diploma requirement).

The only kind of cases in which high school degree requirements tend to be upheld are those involving police department selection criteria. See, e.g., *Castro v. Beecher*, 459 F.2d 725, 735 (1st Cir. 1972); *League of United Latin Am. Citizens v. City of Santa Ana*, 410 F. Supp. 873, 901-02 (C.D. Cal. 1976) (upheld for police, struck down for fire department); *Arnold v. Ballard*, 390 F. Supp. 723, 728 (N.D. Ohio 1975).

⁶³ See, e.g., *James v. Stockham Valves & Fittings Co.*, 559 F.2d 310, 344, 355 (5th Cir. 1977), *cert. denied*, 434 U.S. 1034 (1978); *League of United Latin Am. Citizens v. City of Santa Ana*, 410 F. Supp. 873, 898 (C.D. Cal. 1976); Comment, *supra* note 24, at 918-20 (strict business necessity test generally applied to nonscored objective criteria).

⁶⁴ 433 U.S. 321 (1977).

⁶⁵ 530 F.2d 1159, 1179 (5th Cir.), *cert. denied*, 429 U.S. 861 (1976).

⁶⁶ *Id.* at 1179 (quoting opinion below, 6 Fair Empl. Prac. Cas. (BNA) 511, 523 (S.D. Ala. 1973)).

fied by business necessity for jobs in both categories. It noted that many of the jobs were extremely dangerous, that poor performance could cause great economic loss, and that studies by a special industry task force had recommended a high school diploma requirement.⁶⁷ The Fifth Circuit, however, held the requirement not justified by business necessity. The court reasoned that the employer could instead provide on-the-job training and use tests designed specifically to aid in promotion decisions.⁶⁸

Courts have been similarly strict in analyzing requirements of relevant experience. In *Crockett v. Green*,⁶⁹ the court struck down a requirement of related job experience for skilled craft workers. Even assuming the validity of the requirement, the court held, it could not be justified as a business necessity, for the employer could always provide workers with on-the-job training.⁷⁰

In upper level job cases involving educational and experience qualifications with a disparate impact, courts have taken a very different approach.⁷¹ Such qualifications have been generally upheld on the basis of defenses rejected out of hand in lower level cases. Courts have been persuaded by evidence of apparent rationality and good faith, such as evidence that a challenged credential is commonly used by comparable em-

⁶⁷ 6 Fair Empl. Prac. Cas. (BNA) at 523.

⁶⁸ 530 F.2d at 1179-82.

⁶⁹ 388 F. Supp. 912 (E.D. Wis. 1975), *aff'd*, 534 F.2d 715 (7th Cir. 1976).

⁷⁰ *Id.* at 919-21; *see also* United States v. San Diego County, 21 Fair Empl. Prac. Cas. (BNA) 402 (S.D. Cal. 1979) (enjoining application to women candidates of a four-year in-service experience prerequisite for promotion to police sergeant, on ground that defendants had not demonstrated the requirement's job-relatedness).

Until the Supreme Court's decision in *International Bhd. of Teamsters v. United States*, 431 U.S. 324 (1977), the federal courts had generally struck down use of seniority systems on *Griggs*-type grounds, ruling that such systems had an unjustified impact on blacks, who tended to be at the bottom of seniority ladders by virtue of previous discrimination. *See id.* at 378-79 & nn.2-3 (Marshall, J., dissenting). The courts refused to find such systems justified by business necessity even when employers claimed that an orderly progression up the seniority ladder was essential to gain the experience necessary for effective performance. Instead, employers were allowed only the limited opportunity to prove that experience of a particular kind was essential for particular jobs. While in *Teamsters* the Court immunized seniority systems from wholesale attack, it did so solely on the basis of the language and legislative history of § 703(h) of title VII, 42 U.S.C. § 2000e-2(h) (1976), which specifically refers to seniority systems. 431 U.S. at 348 & n.30, 349-50, 352-54. The Court did not argue that seniority systems could be justified on the basis of business necessity, nor did it otherwise undermine the cases holding that the use of experience as a qualification — other than as part of a seniority system — was subject to the *Griggs* doctrine.

⁷¹ *See generally* B. SCHLEI & P. GROSSMAN, *supra* note 12, at 40-41 (Supp. 1979) (objective nonscored criteria tend to be struck down in lower level and upheld in upper level job cases).

ployers.⁷² Courts seem to be applying a "legitimate business purposes" test — the test rejected in *Griggs* in favor of the business necessity test.

A Second Circuit case, *Townsend v. Nassau County Medical Center*, illustrates the extent to which courts have reversed the *Griggs* presumption.⁷³ The plaintiff, a black technician in a blood bank, had been fired after many years' service because she had failed a newly imposed exam requirement and lacked the bachelor of science degree required to take it again. She had then been rehired to perform exactly the same job in a lower civil service classification at less pay. District Judge Weinstein had ruled that the degree requirement violated title VII because it had a disparate impact on blacks, the defendant had not established that it was job related, and the plaintiff had demonstrated her actual qualification through on-the-job performance.⁷⁴ On appeal the Second Circuit rejected plaintiff's disparate impact claim, arguing that Judge Weinstein had erred in relying on general population statistics to assess the impact of a degree requirement on blacks.⁷⁵ The Second Cir-

⁷² Several cases demonstrate the courts' application of this significantly softer standard to upper level educational prerequisites. In *Campbell v. Ramsay*, 484 F. Supp. 190 (E.D. Ark.), *aff'd per curiam*, 631 F.2d 597 (8th Cir. 1980), the court upheld a university's failure to renew a female plaintiff's teaching appointment in a mathematics department; the university had based its decision on the plaintiff's lack of a Ph.D. The court conceded that the plaintiff "was an excellent teacher" and that "a Ph.D. could not have taught her courses any better than she did." *Id.* at 195. But the court was persuaded by arguments that the Ph.D. requirement served the university's goal of increasing its prestige and its drawing power. In *Scott v. University of Del.*, 455 F. Supp. 1102 (D. Del. 1978), *aff'd in relevant part*, 601 F.2d 76 (3d Cir.), *cert. denied*, 444 U.S. 931 (1979), the court upheld a Ph.D. requirement for promotion in a sociology department despite "surprisingly sparse" evidence of justification. *Id.* at 1124. In *Keyes v. Lenoir Rhyne College*, 15 Fair Empl. Prac. Cas. (BNA) 914 (W.D.N.C. 1976), *aff'd*, 552 F.2d 579 (4th Cir.), *cert. denied*, 434 U.S. 904 (1977), the court upheld a doctorate requirement on the ground that a Ph.D. degree "presumably is evidence of certain skills and knowledge from which one might predict job performance and is an educational credential which is almost universally required for higher ranks in most colleges." *Id.* at 923. In *Wade v. Mississippi Coop. Extension Serv.*, 372 F. Supp. 126 (N.D. Miss. 1974), *aff'd in relevant part*, 528 F.2d 508 (5th Cir. 1976), the court upheld the use of master's and doctor's degrees as factors in filling various professional positions in the state's educational extension services agency for agriculture and home economics; the court relied in part on the national trend to demand advanced degrees for extension programs. *Id.* at 134.

For an upper level case upholding experience on the ground that it was "useful," see *EEOC v. New York Broadcasting Serv.*, 542 F.2d 356 (6th Cir. 1976) (employer's use of experience as a basis for pay differentials upheld despite evidence that women had been deprived of such experience by discrimination).

⁷³ *Townsend v. Nassau County Medical Center*, 558 F.2d 117, 120 (2d Cir. 1977), *cert. denied*, 434 U.S. 1015 (1978).

⁷⁴ See *id.* at 119 (citing unpublished district court opinion).

⁷⁵ *Id.* at 120.

cuit's view is at odds with traditional title VII impact analysis.⁷⁶ The court's apparent concern was with the radical consequences of forcing employers to defend advanced educational degree requirements:

If we were to hold that a bare census statistic concerning the number of blacks in the *general population* who have college degrees could establish a *prima facie* case of discrimination, every employer with a college degree requirement would have the burden of justifying the degree requirement as job-related The requirement of a college degree, particularly in the sciences, seems to be in the modern day of advanced scientific method, a neutral requirement for the protection of the public. . . . There will be time, if a showing of racial impact is made, for the comparison of the requirement of a degree in medicine, law, engineering or other professions with such a requirement for a laboratory technologist⁷⁷

The contrast to *Griggs* and its progeny could hardly be more stark. There the courts found newly imposed qualifications suspect; they accepted as evidence that such qualifications were *not* essential the fact that employees hired earlier did not possess them and yet managed nonetheless to do a satisfactory job.⁷⁸ But in *Townsend* the court upheld an educational qualification as applied to an individual the court conceded to be qualified. In *Griggs* the Supreme Court had condemned the diploma requirement in language suggesting a general skepticism about the need for traditional educational credentials: "History is filled with examples of men and women who rendered highly effective performance without the conventional badges of accomplishment in terms of certificates, diplomas, or degrees."⁷⁹ In *Townsend* the Second Circuit distorted impact analysis, expressing a deep concern that the use

⁷⁶ When an employer sets an absolute prerequisite to employment, such as a degree requirement, general population statistics indicating the relative number of blacks and whites with such degrees are ordinarily the appropriate statistics for assessing the discriminatory impact of the employer's policy. Statistics relating only to applicants for the employment at issue — which the Second Circuit in *Townsend* implied would have been more appropriate — are likely to be irrelevant, for a person who does not possess a degree that is a prerequisite to employment is not likely to apply. In *Griggs*, 401 U.S. at 430 n.6, and *Dothard*, 433 U.S. at 329–31, the Supreme Court relied on general population statistics in dealing with absolute prerequisites to employment; in *Dothard* the Court discussed the reason why such statistics were appropriate in these kinds of cases. *Id.* at 330.

⁷⁷ *Townsend*, 558 F.2d at 120 (citations omitted) (emphasis in original).

⁷⁸ *Griggs*, 401 U.S. at 430 n.6.

⁷⁹ *Id.* at 433.

of educational degrees as prerequisites to various professional endeavors be protected from legal scrutiny.

(b) *Qualifications Assumed Valid in Analyzing Racial Impact.* — Identification of the appropriate candidate pool and its racial makeup is usually the starting point for impact analysis. The *Griggs* test is triggered when a selection device selects minorities in a proportion smaller than that in the pool of qualified candidates.

In lower level cases, courts have accepted rough approximations of the qualified pool. In the case of entry level jobs, for example, courts often look to the racial composition of the general population in order to assess the racial impact of employers' selection procedures. Courts have assumed that persons in the general population are qualified or could easily become qualified. Similarly, courts have used the general work force as an appropriate pool in order to determine the racial impact of procedures for selecting supervisors. When employers have argued that special qualifications are required for a particular job and that the appropriate pool was therefore smaller and contained a significantly higher proportion of whites, courts have required employers to demonstrate the business necessity of such qualifications. These courts were presented with no evidence that all persons in a given pool were qualified for the skilled and responsible jobs such cases often involved. Rather, in order to analyze the racial impact of the employer's policies, pools were identified on the basis of generous assumptions about the number of people with conceded minimal qualifications for the job. For example, in *Grant v. Bethlehem Steel Corp.*,⁸⁰ the employer argued that plaintiffs could not establish disparate impact on the basis of general work force statistics, because experience as an ironworker or as a foreman was clearly a relevant qualification for work as a foreman responsible for the lives of high-construction steelworkers. The Second Circuit held that the defendant had to demonstrate the business necessity of such qualifications; a *Griggs* challenge could not be thwarted by limiting the qualified pool based on casual assumptions that particular qualifications held disproportionately by whites were necessary for adequate job performance.⁸¹

⁸⁰ 635 F.2d 1007 (2d Cir. 1980), *cert. denied*, 101 S. Ct. 3083 (1981), *discussed at supra* pp. 953-54.

⁸¹ *Id.* at 1019. Similarly in *Swint v. Pullman-Standard*, 539 F.2d 77 (5th Cir. 1976), *cert. granted*, 451 U.S. 906 (1981) (No. 80-1190), the court stuck down a promotional system for jobs involving the supervision of welding and steel construction work. The court below had found general work force statistics of limited relevance, because more blacks than whites were functionally illiterate and because black em-

By contrast, in upper level cases courts often simply take for granted the business necessity of education or experience requirements that shrink and "whiten" the pool.⁸² By doing so, courts deny plaintiffs the chance to challenge those qualifications, and free employers of the burden of defending them.⁸³

ployees had fewer relevant skills and less experience than whites. See *id.* at 104. Rejecting this analysis, the Fifth Circuit held that both educational and experience qualifications would have to be justified by the employer on business necessity grounds. *Id.* For similar treatment of these issues, see, e.g., *Watkins v. Scott Paper Co.*, 530 F.2d 1159, 1192-94 (5th Cir.), *cert. denied*, 429 U.S. 861 (1976); *Rowe v. General Motors Corp.*, 457 F.2d 348, 358 (5th Cir. 1972).

⁸² In *EEOC v. E.I. duPont de Nemours & Co.*, 445 F. Supp. 223 (D. Del. 1978), the court found no disparate impact in analyzing defendants' "above career level" promotional pattern. The court assumed the relevance of years of experience as a qualification. In *Agarwal v. Arthur G. McKee & Co.*, 16 Empl. Prac. Dec. (CCH) ¶ 8301 (N.D. Cal. 1977), *aff'd*, 644 F.2d 803 (9th Cir. 1981), the court rejected a racial challenge to the selection and promotion system on the ground that no impact was shown. It implicitly rejected use of the general employer work force as the relevant pool, although the employer had no minimum educational prerequisites and had a policy of filling higher level positions from within by promotion or transfer. *Id.* at 5578-79; see also *Wheeler v. Armco Steel Corp.*, 471 F. Supp. 1050 (S.D. Tex. 1979) (accepting, without discussion, women's lack of relevant prior experience in previously male-dominated world of oil-drilling equipment distribution as adequate explanation for lack of women in top management).

For an upper level job case taking a different approach, see *Greenspan v. Automobile Club*, 22 Empl. Prac. Dec. (CCH) ¶ 30,812 (E.D. Mich. 1980), a sex discrimination case involving professional, technical, and sales positions. For discussions of these general issues, see Copus, *The Numbers Game Is The Only Game In Town*, 20 How. L.J. 374, 382-90 (1977); Shoben, *Probing the Discriminatory Effects of Employee Selection Procedures with Disparate Impact Analysis Under Title VII*, 56 TEX. L. REV. 1, 13-19 (1977).

⁸³ See generally Copus, *supra* note 82, at 390-92 (a prima facie showing of discrimination is made when a substantial disparity exists between the racial composition of the employer's work force and that of the general population); Shoben, *supra* note 82, at 15-16 (courts should resolve doubts in favor of a general population comparison rather than a relevant labor market approach, lest the relevant labor market be too narrowly defined).

An analogous problem arises when courts dismiss disparate treatment cases on the ground that plaintiffs have failed to account for every possible factor, subjective or objective, that might conceivably explain the differential treatment. In *Agarwal*, 16 Empl. Prac. Dec. (CCH) ¶ 8301 (N.D. Cal. 1977), *aff'd*, 644 F.2d 803 (9th Cir. 1981), minority plaintiffs challenged employer practices preventing them from reaching top salary levels, and introduced a multiple regression study controlling for numerous factors including years of education, years since receipt of highest degree, age of employee, type of professional registration held by employee, and years of experience. The court found plaintiffs' regression study inadequate because: (1) the study excluded information that "could have had some bearing upon salary"; (2) the study did not control for type or quality of prior experience and education; and (3) "plaintiff did not attempt to determine whether particular individuals had special abilities or characteristics which might have had a bearing upon salary level." *Id.* at 5574. A similar case is *Presseisen v. Swarthmore College*, 442 F. Supp. 593, 616 (E.D. Pa. 1977),

A similar problem arises when the courts define the qualified pool by looking to the current work force in the relevant industry or to other figures that may simply reflect the kind of discrimination being challenged. For example, to determine whether the racial composition of a particular bank's work force indicates that the bank's selection practices have a disparate impact, a court might take the relevant pool to be the work force in the banking industry as a whole.⁸⁴ The problem is that, if plaintiffs are challenging selection procedures that

aff'd mem., 582 F.2d 1275 (3d Cir. 1978). *Agarwal* and *Presseisen* are discussed and criticized in Finkelstein, *The Judicial Reception of Multiple Regression Studies in Race and Sex Discrimination Cases*, 80 COLUM. L. REV. 737, 741-45 (1980).

Requiring plaintiffs in disparate treatment cases to account for every factor that might explain differential racial treatment places an impossible burden on plaintiffs, who are much worse equipped than employers to know how qualifications are actually assessed. The requirement is also inconsistent with the logic of the disparate treatment doctrine. At the initial stage of disparate treatment cases, plaintiffs have traditionally not been required to prove that qualifications are *in fact* comparable, but simply that they *appear* sufficiently comparable to warrant an explanation from the employer for the disparate treatment. The purpose of the doctrine is to force the employer, and not the plaintiff, to produce that explanation. For example, in *James v. Stockham Valves & Fittings Co.*, 559 F.2d 310 (5th Cir. 1977), *rev'd* 394 F. Supp. 434 (N.D. Ala. 1975), *cert. denied*, 434 U.S. 1034 (1978), a simple showing that blacks were assigned to the worst jobs, that they received less pay even after accounting for seniority, and that there were virtually no objective standards at work was sufficient to make out a *prima facie* case. *Id.* at 328-31; see *Davis v. Califano*, 613 F.2d 957, 964 (D.C. Cir. 1979) (adopting similar analysis in an upper level case). See generally *D. BALDUS & J. COLE, STATISTICAL PROOF OF DISCRIMINATION* § 6.222[13], at 191-94 (1980) (if relevance of a qualification in defendant's selection process cannot be decided on threshold inquiry, plaintiff should not be required to account for it in her *prima facie* case).

⁸⁴ In *Presseisen v. Swarthmore College*, 442 F. Supp. 593 (E.D. Pa. 1977), *aff'd mem.*, 582 F.2d 1275 (3d Cir. 1978), a sex discrimination case, the court determined the relevant labor pool, in part, by looking to the faculties of other educational institutions in the United States. *Id.* at 621; see also *Agarwal v. Arthur G. McKee & Co.*, 16 Empl. Prac. Dec. (CCH) ¶ 8301, at 5576-77 (N.D. Cal. 1977) (using general national and industrial labor market figures for positions comparable to those at issue in the case), *aff'd*, 644 F.2d 803 (9th Cir. 1981); *Crocker v. Boeing Co.*, 437 F. Supp. 1138, 1184 (E.D. Pa. 1977) (in assessing selection procedures for managerial employees, court looked to census figures showing the percentage of blacks employed in related occupational categories).

In *Vuyanich v. Republic Nat'l Bank*, 505 F. Supp. 224 (N.D. Tex. 1980), the court accepted the defendant's argument that, because the bank's upper level employees tended to have certain degrees, the court should look to relevant work force statistics indicating the ethnic makeup of the group possessing such degrees in order to assess the legality of the bank's selection procedures. Because plaintiffs' challenge was directed in part to the legality of the criteria used by the bank in making selection decisions, including bank reliance on just such degrees, the court's approach effectively denied plaintiffs' right to present their discrimination claim.

are common in an industry, it makes no sense to take as the candidate pool a group that has already been screened by such procedures.

Appropriate pools for impact analysis will often be more limited for upper level jobs, for there will often be little dispute that the jobs at issue require some educational and other qualifications beyond those possessed by the general population. There are conceded minimum qualifications for many of these jobs, and these qualifications must be used to define the qualified pool. But the courts' approach in many upper level cases is so strict that it prevents the possibility of challenge to policies going beyond conceded minimum qualifications.⁸⁵

2. *Subjective Selection Systems.* — The significance of title VII on the upper level depends largely on how carefully the courts scrutinize subjective evaluation processes.⁸⁶ Subjective assessments play a role in most upper level employment decisions regarding hiring, promotion, job placement, and salary. Tests and objective criteria such as education and experience requirements are ordinarily used on the upper level primarily as minimum qualifications for certain positions. Once minimum qualifications are met, they and other objective criteria are usually considered only as part of an overall subjective assessment, which is typically based on a variety of subjective procedures: an interview, an evaluation of biographical information, an evaluation of performance in previous educational or work settings.⁸⁷

At the lower level, subjective processes that have an ad-

⁸⁵ Some confusion may result from a misreading of the Supreme Court's decision in *Hazelwood School Dist. v. United States*, 433 U.S. 299 (1977). There the Court distinguished between cases in which general population statistics would be relevant because the skill was one that many "possess or can fairly readily acquire" and those in which "special qualifications are required"; in the latter, general population comparisons "may" have little probative value. *Id.* at 308 n.13. However, there was no dispute in *Hazelwood* about whether teachers constituted the relevant pool. Moreover, the *Hazelwood* Court reaffirmed the principle that the qualified pool should be defined to include those capable of becoming qualified with relative ease. *Id.*

⁸⁶ Authorities dealing with the use of subjective processes in upper level employment include Hunt & Pazuniak, *supra* note 6; Newman, *Remedies for Discrimination in Supervisory and Managerial Jobs*, 13 HARV. C.R.-C.L. L. REV. 633 (1978); Stacy, *Subjective Criteria in Employment Decisions Under Title VII*, 10 GA. L. REV. 737 (1976); Comment, *supra* note 6.

⁸⁷ The assessment center techniques that have become increasingly popular devices for selecting managerial level employees are essentially sophisticated subjective selection processes. Assessment centers attempt to measure an individual's managerial ability through the use of such devices as business games, group discussion problems, in-basket tests, motivational and personality tests, and interviews. See D. BRAY, R. CAMPBELL & D. GRANT, *FORMATIVE YEARS IN BUSINESS: A LONG-TERM AT&T STUDY OF MANAGERIAL LIVES* (1974).

verse racial impact have almost uniformly been condemned. The courts have generally relied at least in part on the theory that subjective processes lend themselves to the expression of conscious or unconscious bias. The oft-cited Fifth Circuit decision in *Rowe v. General Motors Corp.*⁸⁸ states:

[P]rocedures which depend almost entirely upon the subjective evaluation and favorable recommendation of the immediate foreman are a ready mechanism for discrimination against Blacks. . . . We and others have expressed a skepticism that Black persons dependent directly on decisive recommendations from Whites can expect nondiscriminatory action.⁸⁹

Following *Rowe's* lead, courts have often found suspect such features as the absence or ineffectiveness of guidelines limiting discretion; the predominance of whites in the group charged with exercising discretion; and the absence of procedures, such as open job posting and built-in review of decisions, designed to guard against bias.⁹⁰ Additionally, courts have often relied on *Griggs* and found that the subjective systems at issue were not justified as job-related.⁹¹

The Supreme Court indicated similar suspicion about subjective processes in *Albemarle Paper Co. v. Moody*.⁹² The Court criticized a performance rating system for its failure to provide adequate guidance to those evaluating performance. Condemning the criterion of success used — “who is doing the better job” — as “extremely vague and fatally open to divergent interpretations,”⁹³ the Court complained that there was

⁸⁸ 457 F.2d 348 (5th Cir. 1972).

⁸⁹ *Id.* at 359 (citations omitted).

⁹⁰ *Id.* at 358–59.

⁹¹ For cases condemning subjective systems at the lower level on some combination of these theories, see *Crawford v. Western Elec. Co.*, 614 F.2d 1300, 1315–17 & nn. 29–30 (5th Cir. 1980); *James v. Stockham Valves & Fittings Co.*, 559 F.2d 310, 328 (5th Cir. 1977), *cert. denied*, 434 U.S. 1034 (1978); *United States v. Hazelwood School Dist.*, 534 F.2d 805, 812–13 (8th Cir. 1976), *rev'd on other grounds*, 433 U.S. 299 (1977); *Senter v. General Motors Corp.*, 532 F.2d 511, 528–30 (6th Cir.), *cert. denied*, 429 U.S. 870 (1976); *Brown v. Gaston County Dyeing Mach. Co.*, 457 F.2d 1377, 1382 (4th Cir.), *cert. denied*, 409 U.S. 982 (1972); *Hester v. Southern Ry.*, 349 F. Supp. 812, 817 (N.D. Ga. 1972), *rev'd on other grounds*, 497 F.2d 1374, 1381 (5th Cir. 1974). For a discussion of the courts' general presumption against subjective employment practices for jobs ranging up to low level supervisory positions, see Newman, *supra* note 86.

⁹² 422 U.S. 405 (1975). Subjective supervisory ratings were at issue because the employer's validity study relied on such ratings in an attempt to demonstrate the validity of the selection tests challenged by plaintiffs. The Court's criticism of the subjective rating system in this context, however, seems equally applicable when the subjective process is itself the subject of the initial challenge.

⁹³ *Id.* at 433.

"no way of knowing precisely what criteria of job performance the supervisors were considering, whether each of the supervisors was considering the same criteria or whether, indeed, any of the supervisors actually applied a focused and stable body of criteria of any kind."⁹⁴ The Court concluded that there was "simply no way to determine whether the criteria *actually* considered were sufficiently related to the Company's legitimate interest in job-specific ability to justify a testing system with a racially discriminatory impact."⁹⁵

Few lower level job cases have considered exactly what would be required to justify a subjective system; the systems tend simply to be struck down outright. Some courts have implied that subjective systems might be saved by extensive procedural reform — development of detailed guidelines, review procedures, and the like.⁹⁶ Others have ordered such reforms.⁹⁷ Still others have insisted on a demonstration of job-relatedness and business necessity.⁹⁸ And some courts have required a demonstration of validity.⁹⁹

What is clear is that subjective systems are judged by a strict standard. In *Rowe* and other blue collar cases, supervisors selected employees for promotion to foreman on the basis of such apparently reasonable criteria as leadership ability, quality of performance on the job, and ability to take orders.¹⁰⁰ Yet the courts condemned the vagueness of, for example, the quality-of-performance criterion, even though quality of performance in a closely related job is one of the best means of predicting future performance¹⁰¹ and even though supervisors are obviously in a good position to assess

⁹⁴ *Id.* (footnote omitted).

⁹⁵ *Id.* (emphasis in original).

⁹⁶ See, e.g., *Hester v. Southern Ry.*, 349 F. Supp. 812, 817 (N.D. Ga. 1972), *rev'd on other grounds*, 497 F.2d 1374 (5th Cir. 1974).

⁹⁷ See, e.g., *Miller v. Continental Can Co.*, 13 Fair Empl. Prac. Cas. (BNA) 1585, 1602-03 (S.D. Ga. 1976) (defendant ordered to develop written standards establishing reasonably objective criteria and to post vacancies).

⁹⁸ See, e.g., *Rogers v. International Paper Co.*, 510 F.2d 1340, 1350-51 (8th Cir.), *vacated on other grounds and remanded*, 423 U.S. 809 (1975).

⁹⁹ See, e.g., *Brito v. Zia Co.*, 5 Fair Empl. Prac. Cas. (BNA) 1203, 1205 (D.N.M. 1972), *aff'd*, 478 F.2d 1200 (10th Cir. 1973).

¹⁰⁰ See, e.g., *Parson v. Kaiser Aluminum & Chem. Corp.*, 575 F.2d 1374, 1383 & nn.21-22, 1386 nn.25-26 (5th Cir. 1978), *cert. denied*, 441 U.S. 968 (1979); *Watkins v. Scott Paper Co.*, 530 F.2d 1159 (5th Cir.), *cert. denied*, 429 U.S. 861 (1976); *Rowe v. General Motors Corp.*, 457 F.2d 348, 353 (5th Cir. 1972).

¹⁰¹ See, e.g., R. GUION, *supra* note 27, at 380; Owens, *Background Data*, in HANDBOOK OF INDUSTRIAL AND ORGANIZATIONAL PSYCHOLOGY 609, 625 (M. Dunnette ed. 1976) ("One of our most basic measurement axioms holds that the best predictor of what a man *will do* in the future is what he *has done* in the past.") [hereinafter cited as HANDBOOK].

quality of performance. Moreover, basing promotions on prior performance gives employees an incentive to perform well. Finally, this kind of system is simple and inexpensive to administer. The pervasive use of such schemes in industrial workplaces suggests that employers thought they served their business purposes.¹⁰²

On the upper level, courts have applied a far more lenient standard. Often they simply assert that subjective decision-making is appropriate, and that is the end of the matter: an employer has no burden to demonstrate the job-relatedness, validity, or business necessity of any particular subjective system.¹⁰³ In one academic promotion case, the court examined a university's procedures and pronounced them satisfactory according to the following standard: "If the criteria used and the procedures followed were reasonable and rationally related to the decision reached this is about as far as the court can go."¹⁰⁴ This standard is, of course, totally at odds with the *Rowe* line of cases. In another academic case, the Third Circuit upheld a system that represents an extreme example of the uncontrolled and unreviewable discretion condemned on the lower level. Tenure decisions were made on the basis of: "1) teaching effort and effectiveness, 2) scholarly activity, and 3) service to the department, the University, the community."¹⁰⁵ Each member of the department was allowed to use his personal view of each factor's importance.¹⁰⁶ In a

¹⁰² Similarly, courts have condemned the use of performance evaluations as a means of making employment decisions regarding salary, promotion, and layoffs, despite the fact that performance evaluations seem a sensible approach to such decisionmaking. See *James v. Stockham Valves & Fittings Co.*, 559 F.2d 310 (5th Cir. 1977) (salaries and promotions), *cert. denied*, 434 U.S. 1034 (1978); *Brito v. Zia Co.*, 478 F.2d 1200, 1206 (10th Cir. 1973) (layoffs).

¹⁰³ See, e.g., *Tuft v. McDonnell Douglas Corp.*, 581 F.2d 1304 (8th Cir. 1978); *Frausto v. Legal Aid Soc'y*, 563 F.2d 1324 (9th Cir. 1977); *EEOC v. E.I. duPont de Nemours & Co.*, 445 F. Supp. 223, 254-55 (D. Del. 1978); *Agarwal v. Arthur G. McKee & Co.*, 16 Empl. Prac. Dec. (CCH) ¶ 8301 (N.D. Cal. 1977), *aff'd*, 644 F.2d 803 (9th Cir. 1981); *Frink v. United States Navy*, 16 Fair Empl. Prac. Cas. (BNA) 67, 69-70 (E.D. Pa. 1977), *aff'd mem.*, 609 F.2d 501 (3d Cir. 1979), *cert. denied*, 445 U.S. 930 (1980); *Keely v. Westinghouse Elec. Corp.*, 404 F. Supp. 573, 579-80 (E.D. Mo. 1975); *Levens v. GSA*, 391 F. Supp. 35 (W.D. Mo. 1975), *aff'd mem.*, 538 F.2d 332 (8th Cir. 1976).

Articles noting that a more lenient standard appears to be applied at the upper level include: Wagner, *Tenure and Promotion in Higher Education in Light of Washington v. Davis*, 24 WAYNE L. REV. 95, 119-21 (1977); Waintroob, *supra* note 6, at 48-62; Comment, *supra* note 6; see also Devine, *supra* note 6, at 429-35 (criticizing use of lenient standard in university hiring and promotion cases).

¹⁰⁴ *Johnson v. University of Pittsburgh*, 435 F. Supp. 1328, 1357 (W.D. Pa. 1977).

¹⁰⁵ *Scott v. University of Del.*, 601 F.2d 76, 79 (3d Cir. 1979), *aff'd in part and vacating in part on other grounds* 455 F. Supp. 1102 (D. Del. 1978), *cert. denied*, 444 U.S. 931 (1979).

¹⁰⁶ *Scott v. University of Del.*, 455 F. Supp. at 1107-08.

sex discrimination case involving selection of an archivist, the court upheld a subjective selection system, arguing that the factors used were "sufficiently capable of objective evaluation as to be unlikely pretexts for discrimination."¹⁰⁷ Without the benefit of guidelines, a male supervisor assessed candidates' "personality," "habits," and ability to plan and implement a project in the absence of detailed supervision.¹⁰⁸

Courts in upper level cases have recognized the concern at the heart of the *Rowe* doctrine — the tendency of insiders to keep out those who are different. Nonetheless, they have been reluctant to take action against subjective systems that permit self-perpetuation. In *Fogg v. New England Telephone & Telegraph Co.*,¹⁰⁹ the court found that the company had advanced a "male-oriented promotional policy"¹¹⁰ by creating two sex-segregated management level job categories. The court nonetheless denied relief, noting plaintiff's "aggressive, ambitious" character:

While these traits are supposedly ones that make for success in business, they also run counter to the tendency of any bureaucratic hierarchy to perpetuate itself and protect its members against any sudden change or disruption of the established routine . . . [W]hat the Company required at this management level was primarily conformity and the ability to get along with other personnel.¹¹¹

The contrast in judicial attitudes toward subjective systems on the two levels is striking. The discrepancy cannot be explained by differences in the subjective systems at issue. Upper level systems typically involve white decisionmakers passing judgment on black candidates. These decisionmakers typically use criteria so vague that they allow the expression of conscious and unconscious bias — the kind of criteria condemned by the *Albemarle* Court as "fatally open to divergent interpretations."¹¹² Upper level systems typically grant decisionmakers significant discretion in determining which candi-

¹⁰⁷ *Adams v. Reed*, 567 F.2d 1283, 1286 (5th Cir. 1978) (footnote omitted).

¹⁰⁸ *Id.* at 1286 & n.8. Indeed, the courts regularly find systems unobjectionable that rely on the kind of vague and general criteria condemned on the lower level. *E.g.*, *EEOC v. E.I. duPont de Nemours & Co.*, 445 F. Supp. 223, 253-54 (D. Del. 1978) (criteria included "initiative," "originality," and "desire to be a supervisor"); *see Thompson v. McDonnell Douglas Corp.*, 416 F. Supp. 972, 977, 982 (E.D. Mo. 1976), *aff'd*, 552 F.2d 220 (8th Cir. 1977).

¹⁰⁹ 346 F. Supp. 645 (D.N.H. 1972).

¹¹⁰ *Id.* at 651.

¹¹¹ *Id.* at 649.

¹¹² *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 433 (1975), *discussed at supra* pp. 974-75; *see, e.g.*, *Yurko*, *supra* note 6, at 476-78 (criteria typically considered in academic decisionmaking).

dates should even be considered for selection. But although these features are sufficiently suspect on the lower level to call for condemnation of the subjective process, they are viewed with benign approval on the upper level.

3. *Tests.* — At the lower level, courts have imposed a heavy burden of justification on employers who use standardized tests to screen employees. Courts have demanded that employers produce evidence of validation, as well as proof of business necessity.¹¹³ By contrast, in some of the few upper level cases involving tests designed specifically to measure job qualifications, courts have required very little in the way of employer justification. In a case challenging a state bar examination, the court granted defendant's motion for summary judgment, stating baldly: "[T]he principles of test validation developed under Title VII do not apply to professional licensing examinations."¹¹⁴ In a case summarily affirmed by the Supreme Court, another district court upheld use of the National Teacher Examinations (NTE) for hiring and setting salaries.¹¹⁵ As Justices White and Brennan noted in dissent, the "validation study" deemed sufficient by the district court failed to show job-relatedness in traditional terms.¹¹⁶ The district court had conceded the absence of evidence showing a correlation between NTE scores and teacher effectiveness, but had excused the deficiency on the ground that "there is, as yet, no satisfactory measure of teaching effectiveness."¹¹⁷

III. THE CASE FOR A SINGLE STANDARD

A. Elitist Nature of the Current Differential Standard

The current differential standard is elitist. The courts distinguish between selection systems primarily on the basis of

¹¹³ While the Supreme Court arguably indicated doubt about some of these principles in *Washington v. Davis*, 426 U.S. 229 (1976), courts have not generally read that case to significantly reduce the burden of justification imposed on lower level employers in title VII cases. See, e.g., *Guardians Ass'n of N.Y. City Police Dep't, Inc. v. Civil Serv. Comm'n*, 633 F.2d 232, 245-47 (2d Cir. 1980), *cert. granted*, 50 U.S.L.W. 3527 (U.S. Jan. 12, 1982) (No. 81-431); *Davis v. County of Los Angeles*, 566 F.2d 1334, 1338-40 (9th Cir. 1977), *vacated as moot*, 440 U.S. 625 (1979); *United States v. City of Chicago*, 549 F.2d 415, 431-32 (7th Cir.), *cert. denied*, 434 U.S. 875 (1977).

¹¹⁴ *Woodard v. Virginia Bd. of Bar Examiners*, 420 F. Supp. 211, 214 (E.D. Va. 1976), *aff'd per curiam*, 598 F.2d 1345 (4th Cir. 1979).

¹¹⁵ *United States v. South Carolina*, 445 F. Supp. 1094 (D.S.C. 1977), *aff'd mem. sub nom. National Educ. Ass'n v. South Carolina*, 434 U.S. 1026 (1978).

¹¹⁶ See *National Educ. Ass'n v. South Carolina*, 434 U.S. at 1027-28 (White, J., dissenting).

¹¹⁷ *United States v. South Carolina*, 445 F. Supp. at 1108 n.13.

the social and economic status of the jobs involved. They have intervened freely in low-status jobs, even when poor performance in these jobs might have threatened significant economic and safety interests.¹¹⁸ But with high-status jobs, a hands-off attitude has prevailed.

Judges defer to the employers with whom they identify, and they uphold the kinds of selection systems from which they have benefited. When they deal with prestigious jobs, the courts show an appreciation of the apparent rationality of the employment procedures at issue and a respect for the decisionmakers involved that can only be explained by the fact that these cases confront the courts with their own world. Judges have a personal investment in traditional selection procedures on the upper level. By contrast, courts can readily strike down a civil service test or a *Rowe*-like subjective system because, not knowing or caring much about how blue collar workers are chosen or promoted, judges find it easy to focus on the social harm of racial exclusion.

In deferring to upper level employers, courts often profess their lack of expertise. Only the universities, the courts say, are competent to assess teacher qualifications and to design appropriate selection and promotion systems. But courts are surely more qualified to intervene in academic decisions, with which they have some familiarity, than to decide who is qualified to serve in highly skilled blue collar jobs or how they should be chosen. It is the courts' expertise, rather than the

¹¹⁸ Illustrative cases are discussed in Parts I and II. Ordinarily the lives at risk have been those of workers involved in the lawsuits and their coworkers or subordinates, as in cases involving work with potentially explosive equipment or under other dangerous conditions. See, e.g., *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), discussed at *supra* pp. 951-53; *Grant v. Bethlehem Steel Corp.*, 635 F.2d 1007 (2d Cir. 1980), cert. denied, 101 S. Ct. 3083 (1981), discussed at *supra* pp. 953-54; *Watkins v. Scott Paper Co.*, 530 F.2d 1159 (5th Cir. 1976), rev'd in part 6 Fair Empl. Prac. Cas. (BNA) 511, 515 (S.D. Ala. 1973), cert. denied, 429 U.S. 861 (1976), discussed at *supra* pp. 966-67.

In other cases the lives at risk have been those of the public beyond the work force. See, e.g., *United States v. Chesapeake & O. Ry.*, 471 F.2d 582, 588 (4th Cir. 1972) (rejecting business necessity defense based on safety in case involving railway brakeman), cert. denied, 411 U.S. 939 (1973); *United States v. St. Louis-San Francisco Ry.*, 464 F.2d 301, 308 (8th Cir. 1972) (en banc) (same), cert. denied, 409 U.S. 1107 (1973).

A number of courts, however, have applied a relaxed title VII standard to jobs apparently involving extreme safety risks to the general public. See, e.g., *Harriss v. Pan Am. World Airways, Inc.*, 649 F.2d 670, 675-77 (9th Cir. 1980) (sex discrimination case upholding mandatory pregnancy leave for flight attendants); *Boyd v. Ozark Air Lines, Inc.*, 568 F.2d 50, 54 (8th Cir. 1977) (sex discrimination case upholding height qualifications for pilots); *Spurlock v. United Airlines, Inc.*, 475 F.2d 216 (10th Cir. 1972) (upholding system for selection of airline pilots).

lack of it, that makes them reluctant to interfere at the upper level. They know these decisionmakers; they sympathize and identify with their concerns and their use of traditional selection methods. Indeed, courts have occasionally noted that upper level cases strike close to home. In an academic case, one court sympathized with the difficulties of upper level selection as follows:

Of a hypothetical twenty equally brilliant law school graduates in a law office, one is selected to become a partner. Extensive discovery would reveal that the other nineteen were almost equally well qualified. Fifty junior bank officers all aspire to become a vice-president — one is selected. And, of course, *even judges are plagued by the difficulty of decision in selecting law clerks out of the many equally well qualified.*¹¹⁹

Courts' identification with upper level employers also makes them less capable of imagining alternatives to traditional selection systems. In dealing with lower level jobs, the courts have had enough distance to weigh the social cost of racial exclusion against the need for traditional systems. Judges have been able to reject with relative ease employer arguments that things must be as they have always been. If, for example, a police or fire department selection scheme systematically excludes blacks, judges have recognized that there must be alternative schemes that would serve the community's interest in police and fire protection as well or better.

Judges must develop this same analytic distance in looking at upper level selection systems. The *Griggs* doctrine encourages such detachment. By rejecting apparent commonsense as a sufficient defense, it forces courts to analyze their own assumptions. By insisting that employers produce evidentiary justifications for their systems, the doctrine educates courts about the actual strengths and weaknesses of these justifications.

B. The 1972 Amendments to Title VII and Their Legislative History

Not only does the language of title VII fail to draw a distinction based on level of employment, but the legislative history of the 1972 amendments also indicates that no distinction was intended. The history reveals both congressional concern with upper level discrimination and a congressional understanding that title VII would remedy that discrimination.

¹¹⁹ *Faro v. New York Univ.*, 502 F.2d 1229, 1232 (2d Cir. 1974) (emphasis added).

One of the 1972 amendments to title VII extended its coverage to academic institutions.¹²⁰ With that amendment, Congress in effect decided that professorships, for example, should not be treated differently from other jobs. A Senate Report stated:

[I]n the higher-paying and more prestigious positions in the institutions of higher learning, blacks constituted only 2.2% of all positions, most of these being found in all-black or predominantly black institutions. . . . The committee believes that it is essential that these employees be given the *same opportunity* to redress their grievances as are available to other employees in the other sectors of business.¹²¹

The same issue arose when Congress specifically considered and rejected a proposal to exempt other upper level occupations, namely physicians and surgeons. Speaking in opposition to the proposed exemption, Senator Javits said:

One of the things that those discriminated against have resented the most is that they are relegated to the position of the sawers of wood and the drawers of water; that only the blue collar jobs and ditchdigging jobs are reserved for them; and that though they built America, and certainly helped build it enormously in the days of its basic construction, they cannot ascend the higher rungs in professional and other life.

Yet, this amendment would go back beyond decades of struggle and of injustice, and reinstate the possibility of discrimination on grounds of ethnic origin, color, sex, religion — just confined to physicians or surgeons, one of the highest rungs of the ladder that any member of a minority could attain — and thus lock in and fortify the idea that being a doctor or a surgeon is just too good for members of a minority, and that they have to be subject to discrimination in respect of it, and the Federal law will not protect them.¹²²

Congress in 1972 also expanded title VII to cover public employment. The legislative history of that amendment dem-

¹²⁰ The Equal Employment Opportunity Act of 1972, Pub. L. No. 92-261, § 3, 86 Stat. 103, 103-04 (1972) (codified at 42 U.S.C. § 2000e-1 (1976)), abolished the exemption for the employment of individuals engaged in educational activities at non-religious educational institutions.

¹²¹ S. REP. NO. 415, 92d Cong., 1st Sess. 12 (1971) (emphasis added) [hereinafter cited as SENATE REPORT].

¹²² 118 CONG. REC. 3802 (1972) (statement of Sen. Javits).

An amendment proposed in 1972 to exclude public and private hospitals and their professional staffs from title VII coverage was defeated. See 118 CONG. REC. 3799-802 (1972), reprinted in STAFF OF SUBCOMM. ON LABOR, SENATE COMM. ON LABOR AND PUBLIC WELFARE, 92D CONG., 2D SESS., LEGISLATIVE HISTORY OF THE EQUAL EMPLOYMENT OPPORTUNITY ACT OF 1972, at 1455-65 (Comm. Print 1972).

onstrates Congress' concern about the absence of minorities from high positions in the federal bureaucracy; moreover, Congress knew that most white collar federal jobs were filled on the basis of civil service tests of a type traditionally subject to strict *Griggs* analysis.¹²³ Finally, Congress explicitly recognized the importance of increasing minority representation in important state governmental jobs. The House Committee Report took note of the fact that blacks were excluded from many white collar governmental jobs, from city managerial positions, from law enforcement agencies, including prosecutors' offices, and from education. The report concluded that:

The problem of employment discrimination is particularly acute and has the most deleterious effect in these governmental activities which are most visible to the minority communities (notably education, law enforcement, and the administration of justice) with the result that the credibility of the government's claim to represent all the people equally is negated.¹²⁴

When Congress decided in 1972 to cover upper level jobs without in any way distinguishing them from lower level jobs, there was extensive congressional discussion of the fact that the central problem of employment discrimination was one not of overt discrimination, but of covert bias and of systematic exclusion of blacks from certain jobs by facially neutral criteria.¹²⁵ The discussion generally endorsed the then-recent *Griggs* decision and the pre-*Griggs* case law establishing the disparate treatment and disparate impact doctrines.¹²⁶

One cannot assume, however, that Congress endorsed the most radical implications of the *Griggs* impact doctrine. The hard question is what burden should be placed on upper level employers to justify selection systems that have a racial impact. Neither the language nor the legislative history of the

¹²³ See H. REP. NO. 238, 92d Cong., 1st Sess. 23-24 (1971), reprinted in 1972 U.S. CODE CONG. & AD. NEWS 2137, 2158-59 [hereinafter cited as HOUSE REPORT].

¹²⁴ *Id.* at 17, reprinted in 1972 U.S. CODE CONG. & AD. NEWS at 2153; see also SENATE REPORT, *supra* note 121, at 10 (expressing concern with exclusion of minorities from important governmental agencies).

¹²⁵ See *International Bhd. of Teamsters v. United States*, 431 U.S. 324, 391-93 & n.21 (1977) (Marshall, J., concurring in part and dissenting in part). See generally SENATE REPORT, *supra* note 121 (covert discrimination is crucial problem); HOUSE REPORT, *supra* note 123 (same).

¹²⁶ See Meltzer, *supra* note 24, at 429 n.31, 437 & n.69; Rutherglen, *supra* note 52, at 719 & nn.186-87.

Meltzer notes Congress' concern that title VII not be used to impose hiring quotas. He argues that the legislative history provides no resolution of any conflict between this concern and Congress' endorsement of *Griggs*. Meltzer, *supra* note 24, at 436-37.

1972 amendments can alone answer that question.¹²⁷ However, title VII does establish a presumption that upper level selection systems should be treated no differently from lower level systems. The issue addressed in the following Section is whether there is something about upper level job selection systems that demands different legal treatment, despite title VII's apparent single standard.

C. Applicability of Traditional Title VII Standards to Upper Level Selection Systems.

1. *Privacy and Associational Interests.* — One justification offered for a differential standard is that privacy and associational rights on the upper level are entitled to special protection. Courts have expressed concern about interfering with the right of law partners and advertising executives to work with whom they please. Personal relationships are said to be vital to performance on this level.¹²⁸

This argument is no more powerful at the upper level than at the lower. Personal relationships are likely to be extremely important to performance on the lower level in many small enterprises. Personal relationships between blue collar workers dependent on each other for their safety are also important, at least to those workers. Moreover, Congress' decision to extend title VII's coverage to all employers with fifteen or more em-

¹²⁷ Nor are the interpretive guidelines promulgated by the relevant enforcing agencies particularly helpful in determining the burden employers should properly bear. The Uniform Guidelines, *supra* note 13, make no distinctions on the basis of an employee's job. Moreover, the history of the Guidelines indicates that various employer groups, including academic institutions, attempted to obtain significant exemption from coverage. 43 Fed. Reg. 38,290, 38,294 (1978).

The Uniform Guidelines do contain language indicating that in some circumstances employers may have a reduced burden to demonstrate job-relatedness. Uniform Guidelines, *supra* note 13, § 1607.6(B). When an informal or unscored procedure produces an adverse impact, the employer "should eliminate the adverse impact, or modify the procedure to one which is a formal, scored [measure] . . . and then validate . . . , or otherwise justify continued use of the procedure in accord with Federal law." *Id.* § 1607.6(B)(1). When formal, scored procedures produce an adverse impact, the Guidelines indicate that the employer should validate if: (1) there is a large enough sample of persons to achieve a statistically significant study; (2) there is a sufficient range of selection procedure and job performance scores to produce representative results; and (3) adequate measures of job performance can be devised. *Id.* §§ 1607.6(B) to .16(U). If validation is not feasible, "the user should either modify the procedure to eliminate adverse impact or otherwise justify continued use of the procedure in accord with Federal law." *Id.* § 1607.6(B)(2).

¹²⁸ See *Hishon v. King & Spaulding*, 24 Fair Empl. Prac. Cas. (BNA) 1303 (D. Ga. 1980), discussed at *supra* p. 960; *EEOC v. Kallir, Philips, Ross, Inc.*, 420 F. Supp. 919, 926-27 (S.D.N.Y. 1976), *aff'd mem.*, 559 F.2d 1203 (2d Cir.), *cert. denied*, 434 U.S. 920 (1977), discussed at *supra* p. 963.

ployees¹²⁹ was a decision to interfere radically with rights of privacy and association in the workplace. The courts' decision in title VII cases to provide reinstatement and promotion as a matter of right¹³⁰ was a decision to interfere with traditional rights of freedom from unwanted employment relationships.

Courts have also indicated that the lifelong nature of tenure warrants according it special immunity.¹³¹ However, the great bulk of the lower level jobs that have been the focus of title VII enforcement have involved something akin to life tenure. They have generally been covered by civil service laws or collective bargaining agreements, both of which provide effective guarantees against discharge for anything other than egregious misconduct. Against this background Congress amended title VII to cover academic institutions. It knew full well that tenure decisions were at issue, and yet provided no special immunity for them.

2. *Difficulties of Demonstrating Job-Relatedness and Business Necessity.* — More significant are the arguments that traditional title VII standards are inappropriate on the upper level because of special difficulties in demonstrating job-relatedness and business necessity.

(a) *Nature of the Jobs.* — It is often argued that upper level employers should be free from any real burden of demonstrating job-relatedness because defining and measuring effective job performance is simply too difficult.¹³² It is often

¹²⁹ Equal Employment Opportunity Act of 1972, Pub. L. No. 92-261, § 2(4), 86 Stat. 103, 103 (codified at 42 U.S.C. § 2000e(b) (1976)). The Act originally covered all employers with 25 or more employees. Civil Rights Act of 1964, Pub. L. No. 88-352, § 701(b), 78 Stat. 241, 253.

¹³⁰ See *supra* pp. 962-63.

¹³¹ See *Labat v. Board of Higher Educ.*, 401 F. Supp. 753, 756 (S.D.N.Y. 1975).

It is also said that academic freedom would be jeopardized by strict enforcement of title VII standards. In *Gray v. Board of Higher Educ.*, 50 U.S.L.W. 2307 (S.D.N.Y. Nov. 9, 1981), the court relied on considerations of academic freedom in denying plaintiff discovery of how particular defendants voted on the reappointment and tenure decision at issue. The Fifth Circuit rejected a comparable claim in *In re Dinnan* (Blaubergs v. Regents of Univ. Sys.), 661 F.2d 426 (5th Cir. 1981); the court found no threat to academic freedom in permitting access to information vital to the proof of discrimination claims. Indeed, the court found that to deny such access and thereby to provide academic institutions with effective immunity from such claims would threaten "the very core of values that [academic freedom] now protects." *Id.* at 430. It noted that "ideas may be suppressed just as effectively by denying tenure as by prohibiting the teaching of certain courses." *Id.* (emphasis in original).

¹³² See, e.g., *United States v. South Carolina*, 445 F. Supp. 1094, 1108 n.13 (D.S.C. 1977) (court excused absence of evidence that success on National Teacher Examinations correlated with success on job, on ground that "there is, as yet, no satisfactory measure of teaching effectiveness"), *aff'd mem. sub nom.* *National Educ. Ass'n v. South Carolina*, 434 U.S. 1026 (1978); *Ash & Kroeker, Personnel Selection*,

difficult to define jobs on the upper level. There is no consensus on what the ideal lawyer or school principal or bank vice-president should be doing. Often there is no single ideal; many different behaviors can be successful. It is also difficult to develop selection devices that predict the complex behavior these jobs require. It would be still more difficult to prove that the devices were selecting those capable of the most effective performance, precisely because we are uncertain about what effective performance is.

Yet these facts cut the other way. Our uncertainty about how certain jobs should be performed and about the usefulness of existing systems for predicting performance makes it hard to justify the continued use of selection devices that effectively exclude blacks. Moreover, we are not certain what ideal job performance means, in part, because definitions of the ideal performance of lawyers or doctors or teachers reflect value-laden choices. These choices are largely political: they involve decisions, for example, about how scarce resources should be allocated and what services should be provided to what social groups. This too argues for including rather than excluding major racial groups in such "political" positions.¹³³

(b) *Nature of the Selection Systems.* — Attempts to justify a differential standard often rely on the difficulties of validating the education and experience requirements and the subjective procedures generally used on the upper level. When looking at comparable selection devices on the lower level, courts have generally found it unnecessary to decide whether validation was required. Such devices tended to be struck down outright, on the ground that they were obviously not justified by business necessity. On the upper level, courts are reluctant to decide that objective qualifications are simply unnecessary or that subjective systems should be outlawed as inherently

Classification, and Placement, 26 ANN. REV. PSYCHOLOGY 481, 493 (1975) (the defense typically raised in cases involving certification examinations is that criteria of success defy measurement); Hunt & Pazuniak, *supra* note 6, at 128 ("[I]t will frequently be impossible to devise reliable measures of job performance or of employee adequacy" for upper level jobs.).

¹³³ Some aspects of upper level jobs may warrant the development of new methods by which employers may satisfy their burden of justification. Such jobs are likely to change over time, thus rendering validation more difficult. And the relatively small number of incumbents and candidates with respect to particular positions may make it impossible to obtain statistically significant results in validation studies under traditional methods of analysis. See R. GUION, *supra* note 27, at 457-76 (discussion of personnel selection at upper levels); see also Trattner & O'Leary, *Sample Sizes for Specified Statistical Power in Testing for Differential Validity*, 65 J. APPLIED PSYCHOLOGY 127 (1980) (small sample sizes make validation difficult). These difficulties simply call for the development of new validation techniques. See *infra* note 174.

suspect. But the obvious alternative is for the courts to demand validation.

Courts often argue that the principles of validation simply do not apply to nonscored objective criteria or that it is inherently impossible to validate these criteria. Hence, the argument goes, commonsense arguments for their use must be accepted. However, the testing experts would counsel that, when such criteria are used as prerequisites to employment, one can systematically examine the connection between the criteria and the abilities essential to adequate performance.¹³⁴ And when such criteria are used as factors in a discretionary process — with extra credit given to those with additional education and experience, for example — validity studies can measure the correlation between additional years of experience and job performance, just as they can measure the correlation between scores on written employment tests and job performance.¹³⁵ Testing experts have in fact made efforts to validate such objective criteria. Although they have produced little evidence demonstrating the validity of employment systems relying on objective criteria, the same was true of scored tests when *Griggs* was decided, and was indeed part of the rationale for the *Griggs* result. It is in part because there is reason to question the validity of current systems relying on objective criteria that a strong burden of justification should be imposed on the employer when such criteria have an adverse racial impact.¹³⁶

¹³⁴ See *infra* p. 1017.

¹³⁵ The Uniform Guidelines make no distinction between nonscored objective criteria and scored tests. Both are included in the definition of "selection procedures," and both must be validated if their use has an adverse racial impact. See Uniform Guidelines, *supra* note 13, §§ 1607.5, .16(Q). See generally E. MCCORMICK & D. ILGEN, *INDUSTRIAL PSYCHOLOGY* 181-209 (7th ed. 1980) (discussing the use of biographical data and interviews in personnel selection); Guion, *Recruiting, Selection, and Job Placement*, in *HANDBOOK*, *supra* note 101, at 777, 799-802 (discussing the need for proof of validity for predictors other than tests, such as biographical data and interviews).

Upper level job cases requiring validation of nonscored objective criteria include: *Greenspan v. Automobile Club*, 495 F. Supp. 1021 (E.D. Mich. 1980) (prior related experience requirements for professional, technical, and sales positions); *Leisner v. New York Tel. Co.*, 358 F. Supp. 359, 368-69 (S.D.N.Y. 1973) (requirements such as prior supervisory experience and technical degrees for management level jobs). In *Townsend v. Nassau County Medical Center*, 11 Empl. Prac. Dec. (CCH) ¶ 10,852 (E.D.N.Y. 1975), *rev'd*, 558 F.2d 117 (2d Cir. 1977), *cert. denied*, 434 U.S. 1015 (1978), Judge Weinstein discussed the applicability of validation techniques to a college degree requirement for a blood bank technician's job. *Id.* at 7567-69.

¹³⁶ See Francis, *Diplomas, Degrees, and Discrimination*, 26 HASTINGS L.J. 1377, 1381 n.23, 1389 n.60 (1975) (arguing that use of credentials by employers typically cannot be justified under *Griggs* analysis and noting evidence that credentials may be inversely related to performance for many jobs); Tenopyr, *The Realities of Employ-*

Some have argued that educational degree requirements are inherently less suspect than the written tests commonly held to require validation, because degrees are conferred on the basis of tests designed not by or for the employer but by independent institutions.¹³⁷ But the chances that a test will prove valid are *enhanced* when it is tailored to predict performance in a particular job. The fact that degrees are granted on the basis of testing processes designed to serve quite different purposes makes their validity as employment criteria more, rather than less, suspect.¹³⁸

Similarly, courts in upper level cases often seem to assume that subjective systems are inherently incapable of being validated.¹³⁹ This assumption is implicit in many of the cases upholding such systems on the ground that some kind of subjective system clearly seems appropriate, without further inquiry into whether the system at issue is valid. Again, there seems to be little basis for the assumption. The major validation methods, discussed in more detail in Part IV, are applicable to all selection devices. The validity of a written test is traditionally established by correlating the scores of candidates who passed the test and are on the job with indicia of their job performance. A subjective assessment process, like a written test, can be scored even if only for the purpose of conducting a validity study. Alternatively, an employer can establish the validity of subjective processes by proving that the behavior assessed by the process represents a fair sample of the behavior required in the job at issue, or that the sub-

ment Testing, 36 AM. PSYCHOLOGY 1120, 1123 (1981) (compared to scored tests, educational requirements have low validity coefficients); White & Francis, *supra* note 30, at 1231-32 (educational degree requirements generally lack job-relatedness and should be subject to *Griggs* standard). *Contra* Lerner, *supra* note 30, at 23-24.

¹³⁷ See *Vuyanich v. Republic Nat'l Bank*, 505 F. Supp. 224, 372-73 & n.195 (N.D. Tex. 1980).

¹³⁸ See generally Francis, *supra* note 136, at 1389-92 (better alternative predictors are usually available); White & Francis, *supra* note 30, at 1243 (characterizing degrees as "cheap approximations of more precise screening devices").

¹³⁹ See, e.g., *Vuyanich v. Republic Nat'l Bank*, 505 F. Supp. 224, 371 (N.D. Tex. 1980); see also *Hunt & Pazuniak*, *supra* note 6, at 128 (validation for upper level subjective evaluation "may not as a practical matter be 'technically feasible'"); Note, *Self Defense for Women Lawyers: Enforcement of Employment Rights*, 4 U. MICH. J.L. REF. 517, 528-29 (1971) (personal interview incapable of validation).

A very few decisions involving upper level jobs have held that employers must validate subjective selection processes. See *Wade v. Mississippi Coop. Extension Serv.*, 372 F. Supp. 126, 142 (N.D. Miss. 1974), *aff'd in part, rev'd in part on other grounds, vacated in part on other grounds, and remanded*, 528 F.2d 508, 518 (5th Cir. 1976); *Leisner v. New York Tel. Co.*, 358 F. Supp. 359, 368-69 (S.D.N.Y. 1973).

jective process measures the knowledge, skills, or abilities that are the necessary prerequisites to the job.¹⁴⁰

The industrial psychology literature does not support the notion that subjective systems should be immune from validation principles. The profession has taken the stand that all selection systems, including subjective ones, can and indeed should be validated.¹⁴¹ The literature contains numerous descriptions of validity studies of the most commonly used subjective processes, such as interviews,¹⁴² the evaluation of biographical data,¹⁴³ and assessment center techniques.¹⁴⁴ The literature does note that the subjective devices most commonly used in upper level employment have rarely been validated, but this observation is thought to argue for the need to do more validation.¹⁴⁵ The literature also reveals little evidence that the subjective processes common to the upper level are valid.¹⁴⁶ Again, exactly the same situation existed with respect

¹⁴⁰ See *supra* pp. 1016-17.

¹⁴¹ "When any selection procedure is used, the essential principle is that evidence be accumulated to show a relationship between decisions based on assessments made by that procedure and criteria such as job performance . . ." DIVISION 14 PRINCIPLES, *supra* note 12, at 1. The Uniform Guidelines are more ambiguous on the issue of the employer's obligation to validate subjective processes. While they provide that, in general, the duty to demonstrate job-relatedness applies to all selection procedures, including subjective ones, they also state that, when "informal or unscored procedures" are used, the employer must either (1) modify the procedure to a formal, scored or quantified one and then validate it, or (2) "otherwise justify [its] continued use . . . in accord with Federal law." Uniform Guidelines, *supra* note 13, § 1607.6(B).

¹⁴² See, e.g., N. MAIER, *PSYCHOLOGY IN INDUSTRIAL ORGANIZATIONS* 226-27 (4th ed. 1973); M. MINER & J. MINER, *EMPLOYEE SELECTION WITHIN THE LAW* 145-65 (1979); Heneman, Schwab, Huett & Ford, *Interviewer Validity as a Function of Interview Structure, Biographical Data, and Interview Order*, 60 J. APPLIED PSYCHOLOGY 748 (1975); Landy, *The Validity of the Interview in Police Officer Selection*, 61 J. APPLIED PSYCHOLOGY 193 (1976).

¹⁴³ See, e.g., M. MINER & J. MINER, *supra* note 142, at 183-201; Dunnette & Borman, *Personnel Selection and Classification Systems*, 30 ANN. REV. PSYCHOLOGY 477, 510 (1979).

¹⁴⁴ See, e.g., Hinrichs, *An Eight-Year Follow-Up of a Management Assessment Center*, 63 J. APPLIED PSYCHOLOGY 596 (1978); Klimoski & Strickland, *Assessment Centers — Valid or Merely Prescient*, 30 PERSONNEL PSYCHOLOGY 353 (1977).

¹⁴⁵ See, e.g., M. MINER & J. MINER, *supra* note 142, at 338; Dunnette, *Personnel Management*, 13 ANN. REV. PSYCHOLOGY 285, 291-92 (1962); Hinrichs, *supra* note 144, at 596.

¹⁴⁶ See generally M. MINER & J. MINER, *supra* note 142 (discussing variety of selection devices commonly used for upper level jobs and summarizing related validity studies); sources cited *supra* notes 132-45 (same).

While claims have been made for the validity of assessment centers, many writers have questioned the significance of the validity studies on which these claims are based. See *infra* pp. 1016-18; see also Hinrichs, *supra* note 144, at 600 (assessment center validity studies "rarely if at all" address the key issue whether the centers succeed in predicting job performance).

to written tests at the time of *Griggs*, and that was thought to provide a reason for developing the *Griggs* requirement of proof of job-relatedness.

The point is not that the validation techniques appropriate on the upper level will be identical to those developed for lower level selection devices. There may well be differences in selection methods that call for development of new techniques. Techniques developed in the context of written tests given in identical form to hundreds of applicants may well be of limited assistance. The point is rather that the fundamental principles of validation apply to all selection devices; acceptance of this concept would lead to the development of techniques appropriate to upper level devices.

3. *Quality of Performance*. — It seems likely that fear of lowering the quality of performance in jobs that are thought to really matter¹⁴⁷ is a major factor underlying development of a differential standard. On the lower level, courts strike down subjective criteria not because exclusively objective criteria would be preferable, but because at that level the possibility that a subjective system might produce better qualified candidates is not thought to be worth the cost of racial exclusion. By contrast, on the upper level courts uphold subjective systems because selection of the best candidates seems essential. Moreover, courts may be reluctant to demand proof of job-relatedness and business necessity on the upper level because such proof will be difficult and expensive to produce. Courts may fear that requiring this proof would seriously impair the quality of performance in two ways. First, the quality of important social institutions might be adversely affected if employers expended resources on validation or related litigation, or if they altered employment devices to suit the views of judges and testing experts. Second, employers may opt for a quota system to avoid the costs of litigation and validation; if we assume that existing employment schemes select the most qualified candidates, we can impose quota systems only at the risk of reducing quality of performance.

A major argument, then, for a relaxed standard on the upper level ultimately rests on a rejection of the meritocratic rationale that has been used to justify the *Griggs* impact doctrine.¹⁴⁸ But if quality of performance is more important on

¹⁴⁷ This concern was expressed by the dissenting judge in *Davis v. Califano*, 613 F.2d 957 (D.C. Cir. 1979): "To my mind this case presents an issue that is of great national concern — is the cry of discrimination going to be used as a means for the promotion of underqualified employees to positions requiring great ability?" *Id.* at 972 (MacKinnon, J., dissenting).

¹⁴⁸ See *supra* pp. 957–59.

the upper level, so also, arguably, is the cost of racial exclusion. Black participation in policymaking jobs at major social institutions is important for a number of reasons, including the potential for affecting the way race-related issues are resolved. Moreover, even if the only concern were quality of performance — if the cost of racial exclusion were ignored — there are strong arguments supporting *Griggs'* meritocratic rationale. It is the constant refrain of civil rights litigants that they are seeking not to destroy but to further merit-based employment. A review of the litigation in this area reveals that there is a good deal to this claim.

In the first place, contrary to the common assumption, the standard imposed in lower level cases has not been impossibly strict. Courts have not blindly rejected all systems not meeting the highest standards of the industrial psychologists. Indeed, in recent years many courts have approved, as adequately job-related, systems based on validity studies falling far short of professional standards.¹⁴⁹

What courts have done is to impose a standard that forced employers to come up with evidence, rather than simply arguments based on commonsense and common practice. Further, the standard was strict enough to force courts to scrutinize closely both the systems at issue and the employers' justifications for those systems. In applying this standard, courts were educated by the testing experts. Courts learned, for example, that certain selection devices that seemed obviously job-related were not. Thus, one study of a high school diploma requirement showed that those with a diploma had a higher turnover rate than those without; high school graduates were more likely to get bored with routine jobs.¹⁵⁰ The courts also learned some fundamental questions to ask in analyzing the job-relatedness of employment systems. Had the employer made any effort to determine what kind of performance was called for in the job at issue? Did the employer have any way of deciding which job incumbents were performing most effectively, and therefore any way of discovering whether its selection devices successfully predicted good performance?

Critics have focused on the technicalities of validation in

¹⁴⁹ See, e.g., *Guardians Ass'n of N.Y. City Police Dep't, Inc. v. Civil Serv. Comm'n*, 630 F.2d 79, 89-90 (2d Cir. 1980) (courts have not interpreted testing theories so strictly as to strike down all tests with disparate impact), *cert. denied*, 101 S. Ct. 3083 (1981).

¹⁵⁰ R. GUION, *supra* note 27, at 491; see also U.S. DEP'T OF HEALTH, EDUC., AND WELFARE, *WORK IN AMERICA* 135-36 (1973) (education and job performance inversely related in many jobs).

claiming that the *Griggs* burden is impossible to meet. But the fact is that employers in the lower level cases lost, for the most part, because they were unable to provide answers to such fundamental questions. They lost because they could produce so little to demonstrate that their existing selection systems were in fact merit-based.

Litigation has also exposed the extent of the antimerit characteristics of traditional employment systems. The systems often seem designed not only to screen out minorities, but also to select, reward, and advance candidates on the basis of factors that have little to do with promoting high quality performance. Employment schemes typically produced by civil service systems and collective bargaining agreements often place a premium on, for example, employee security and the protection of employees against individual favoritism or bias. The concept of reward and advancement on the basis of merit is to a great degree sacrificed in favor of such interests.

When courts have investigated systems traditionally used to select sanitation workers, police officers, firefighters, and school principals, they have found written examinations that were absurd on their face. Exam questions demanded knowledge that a promising candidate had no reason to possess.¹⁵¹

¹⁵¹ For example, the 1961 examination for New York City elementary school principals contained the question:

Of the following characters in the nursery rhyme, THE BURIAL OF POOR COCK ROBIN, the one who killed Cock Robin is the

1. Lark
2. Thrush
3. Bull
4. Sparrow

A question on the 1965 exam for junior high school assistant principals in New York City went as follows:

I've Got a Little List, from the *Mikado* is sung by

1. Nanki-Poo
2. Pish-Tush
3. Ko-Ko
4. Pooh-Bah

Another multiple choice question on this examination asked: "Which one of the following violin makers is *NOT* of the great triumvirate of Cremona?" *Chance v. Board of Examiners*, 330 F. Supp. 203, 220 n.23 (S.D.N.Y. 1971), *aff'd*, 458 F.2d 1167 (2d Cir. 1972). The district court noted: "[M]any of [these] questions strike us as having little relevance to the qualities expected of a school supervisor." *Id.* at 220. The authors served as plaintiffs' counsel in the *Chance* case and in several other cases cited in this Article.

New York City firefighters were tested on their knowledge of such words as "attest," "luminous," "deficit," and "irate." *Vulcan Soc'y of the N.Y. City Fire Dep't v. Civil Serv. Comm'n*, 490 F.2d 387, 397 n.12 (2d Cir. 1973). The court observed: "It is hard to understand how the ability to find the closest analogue to most of these words is a good test of the ability to fight fires" *Id.*; see also *Boston Chapter, NAACP, Inc. v. Beecher*, 504 F.2d 1017, 1022 (1st Cir. 1974) (Boston firefighters were tested on the spelling of such words as "pressurized" and "buoyancy"), *cert. denied*, 421 U.S. 910 (1975).

To succeed on such examinations, candidates took cram courses designed to teach mnemonics. Candidates for the New York City school principal examinations were advised to prepare for essay questions on such topics as "Developing a Learning Experience" by memorizing "RICE" ("Readiness; Instructional materials provided; Carrying out the experience; Evaluating the experience"). They were taught to discuss the topic of "Improving School Discipline" by memorizing "PERT CAGES" ("Planning of standards and rules; Environment, improvement of," and so on).¹⁵²

The courts found that such examinations were used as the almost exclusive selection device; no consideration was given, for example, to personal qualities that might make police officers likely to function well under stress, to abide by rules, or to relate well to the community. In case after case, the courts struck down these systems, not merely because defendants failed to prove that their exams were job related, but because the evidence as a whole demonstrated that the systems were irrational. In *Chance v. Board of Examiners*, the Second Circuit struck down the use of the school principal examination described above.¹⁵³ For decades educators had argued that the rigidities of the selection system and the arbitrariness of the examinations discouraged many of the most qualified candidates from applying. Critics claimed that the scheme perpetuated the mediocre calibre and bureaucratic mentality of the supervisory staff. When suit was brought, the superintendent of the city school system refused to defend; he agreed that the system was an obstacle to improving the quality of supervisory personnel. The district court found the written examination system to be an absurd and irrational method for selecting school supervisors; the Second Circuit suggested that it was no better than drawing names out of a hat.¹⁵⁴

¹⁵² *Chance v. Board of Examiners*, 330 F. Supp. 203, 222 & n.24 (S.D.N.Y. 1971), *aff'd*, 458 F.2d 1167 (2d Cir. 1972).

¹⁵³ *Chance v. Board of Examiners*, 458 F.2d at 1175.

¹⁵⁴ *Id.* In *Association Against Discrimination in Employment v. City*, 25 Fair Empl. Prac. Cas. (BNA) 1013 (2d Cir. 1981), the court reviewed the evidence regarding the challenged firefighters' examination as follows:

[T]he exam had no rational relationship to the skills needed in firefighting. . . . [T]he Bridgeport Fire Chief had testified at the remedy hearing "that there might be an inverse correlation between those who passed the exam and those who are most qualified to be firefighters." For example, while the Fire Chief stated that superior physical ability and intelligence are the two most important attributes, the 1975 hiring process used simply a pass-fail physical agility test, and the 1975 written exam weighed negatively any high scores on comprehension questions. . . . Ironically, the court noted that the 1975 exam "represented a decided improvement over some of the earlier civil service exams employed by Bridgeport."

Id. at 1017 (citations and footnotes omitted).

Civil service schemes for choosing whom to promote often seem similarly at odds with merit. In many cases, little or no credit is given for the quality of prior performance. Thus, the most promising predictor of future performance is ignored.¹⁵⁵ Nor is there much incentive to perform well in entry level jobs, because the quickest route to promotion is not to focus on the immediate job but to study for a written examination by taking cram courses. These examinations, like entry level examinations, have often been condemned by the courts as irrational.¹⁵⁶

Litigation has also revealed the institutional bias that use of these examinations represents. It was known for years that black candidates as a group scored much lower than whites. As a result, disproportionately few blacks made the cutoff score required for eligibility. Moreover, many civil service systems ranked candidates according to their test scores; authorities were generally required to hire from the top of the ranked list. Since blacks who managed to make the cutoff score were usually ranked low on the list, they were the last to be hired. Indeed, if not hired before a new examination was given, they might never be hired, for the old eligibility list was usually eliminated.

The examinations placed a premium on mastery of the English language and on traditional test-taking capacities — abilities often irrelevant to the jobs being allocated. Emphasizing these abilities did, however, put blacks and Hispanics at a significant disadvantage. As a result, minorities were excluded from many jobs for reasons that had little bearing on their likely ability to perform. A vivid illustration was New York City's reliance on a traditional civil service examination to select sanitation workers. The pool of those likely to be interested in these low-skilled jobs was significantly black and Hispanic, and many in the Hispanic group had limited

¹⁵⁵ See sources cited *supra* note 101.

¹⁵⁶ See, e.g., *Firefighters Inst. for Racial Equality v. City of St. Louis*, 549 F.2d 506, 511 (8th Cir.) (written test for promotion to fire captain "failed to test the one major job attribute that separates a firefighter from a fire captain, that of supervisory ability" (emphasis in original)), *cert. denied*, 434 U.S. 819 (1977); *Kirkland v. New York State Dep't of Correctional Servs.*, 374 F. Supp. 1361, 1378 (S.D.N.Y. 1974) (examination for promotion to correction sergeant failed to test essential qualifications: leadership, understanding of inmate resocialization, ability to empathize with persons from different backgrounds, and ability to cope with crises), *aff'd in relevant part*, 520 F.2d 420 (2d Cir. 1975), *cert. denied*, 429 U.S. 823 (1976); see also *Officers for Justice v. Civil Serv. Comm'n*, 371 F. Supp. 1328, 1338 (N.D. Cal. 1973) ("[T]he construction of the examination does not reveal the careful attention to the selection of questions necessary to insure that the examination accurately tested candidates' knowledge, skills, or attitudes.").

English literacy. For New York to rely on such an examination to fill sanitation department jobs — jobs that required minimal written or oral communicative abilities — seemed a fairly blatant form of institutional racism.¹⁵⁷

Litigation has also revealed the extent to which seniority systems have sacrificed merit selection and high quality performance in the name of other interests. Seniority systems generally prevent the rapid promotion of the best qualified workers and thereby limit incentives for good performance on lower levels. They also inject into job placement decisions factors that are completely at odds with business safety and efficiency. Seniority systems normally reward those with the greatest seniority by giving them the most desirable jobs, which are often the more difficult, responsible jobs. Seniority promotion runs contrary to merit if the more senior are no longer as capable as younger workers. On the other hand, the most desirable jobs sometimes involve easier work and less responsibility; a merit system might place the least experienced workers in these jobs.¹⁵⁸

Seniority systems also serve to protect the jobs of employees whom employers might consider high risk, even when those jobs entail significant responsibilities. In a case involving the New York City Transit Authority, an employment policy excluding former drug addicts, even those long engaged in treatment and rehabilitation programs, was defended on grounds of business necessity. The Transit Authority argued that all its jobs were safety-sensitive and that exclusion of those with drug histories was necessary because of their questionable reliability. Discovery revealed that a variety of potentially high-risk groups, such as alcoholics and heart disease victims, received significant protection as a result of seniority and other protective provisions under the collective bargaining agree-

¹⁵⁷ This system was challenged in *Lugo v. Bronstein*, No. 73 Civ. 4886 (S.D.N.Y. filed Nov. 14, 1973). The case was resolved by a settlement accepting a substantially revised examination system requiring only minimal English language and arithmetic skills, and the ability to read street signs. *Lugo v. Bronstein*, No. 73 Civ. 4886 (S.D.N.Y. order accepting settlement and dismissing case June 26, 1974).

¹⁵⁸ In *Hodgson v. Greyhound Lines, Inc.*, 499 F.2d 859 (7th Cir. 1974), cert. denied, 419 U.S. 1122 (1975), the employer successfully defended on grounds of business safety and necessity against an age discrimination claim brought under the Age Discrimination in Employment Act, 29 U.S.C. §§ 621-634 (1976 & Supp. III 1979). Litigation revealed a seniority system that seemed inconsistent with safety goals. The system meant that new and inexperienced drivers were placed in "extra-board" jobs, in which driver accidents were most likely because of the long hours and arduous work. At the same time, older workers, who the company asserted might be incapable of performing extra-board work safely, were allowed by virtue of their seniority to choose such work if they wanted the extra money it provided.

ment. Alcoholics who had worked long enough to gain significant seniority would not be fired even if caught drinking on a conceded safety-sensitive job, such as motorman. Alcoholics could join a Transit-Authority-sponsored alcoholic treatment program while continuing to perform in safety-sensitive positions.¹⁵⁹

Antimeritocratic influences have also been exposed in the subjective systems often used at the lower level. Court scrutiny here has revealed the degree to which job allocation depends on friendship and nepotism¹⁶⁰ and on factors directly linked to social status and race.¹⁶¹

The point is not that traditional employment schemes are inevitably inconsistent with business safety and efficiency.¹⁶² The point is simply that litigation has helped reveal that the schemes have not always been designed primarily with a view to merit selection. Rather, they appear to have been designed to serve a wide variety of interests — many of them in conflict with merit selection. Challenges to such employment schemes as racially exclusionary cannot, therefore, simply be characterized as demands that merit principles be compromised. They can be seen to further those principles by exposing the real workings of job allocation schemes and making it clear which aspects of the systems serve valid interests and which do not.

The judicial record with respect to upper level employment systems is sparse because so few have been subject to traditional title VII scrutiny. But there is no reason to assume that

¹⁵⁹ *Beazer v. New York City Transit Auth.*, 399 F. Supp. 1032 (S.D.N.Y. 1975), *aff'd in part and rev'd in part*, 558 F.2d 97 (2d Cir. 1977), *rev'd*, 440 U.S. 568 (1979).

¹⁶⁰ *See, e.g., Grant v. Bethlehem Steel Corp.*, 635 F.2d 1007 (2d Cir. 1980), *cert. denied*, 101 S. Ct. 3083 (1981). The defendant argued that the dangerous nature of steel construction work justified its subjective system for selecting foremen. *Id.* at 1012, 1016. The court discovered that jobs as foremen were handed out by the white supervisors to their white friends and relatives, often without regard to safety considerations; supervisors had hired persons with drinking problems and persons responsible for prior accidents. *Id.* at 1018-19.

¹⁶¹ *See, e.g., United States v. City of Chicago*, 549 F.2d 415, 432 (7th Cir.) (striking down background investigation procedure involving inquiries into social status and arrest records of a candidate's family members), *cert. denied*, 434 U.S. 875 (1977).

¹⁶² *Cf. Feller, A General Theory of the Collective Bargaining Agreement*, 61 CALIF. L. REV. 663, 768-69 (1973) (collective bargaining agreements further a sense of fairness that is important to the employees' effectiveness and that may, in terms of overall economic efficiency, justify the antimerit aspects of the promotion and lay-off decisionmaking such agreements entail); Freeman & Medoff, *The Two Faces of Unionism*, PUB. INTEREST, Fall 1979, at 69 (discussing conflicting views on whether unionism fosters or negates economic efficiency, and arguing that in many settings unionism has more positive than negative effects).

upper level employment systems are any more pure in their merit orientation than lower level systems. Wealth, social status, and personal relationships, to take a few examples, are of immense importance in determining who gets selected for the most prestigious positions. Many employers favor graduates of institutions from which their own top level employees graduated. Upper level subjective systems typically invest decisionmakers with broad discretion, allowing the expression of personal bias and inviting selection of candidates who resemble those doing the selecting. A major New York law firm described and defended its procedure for selecting associates as follows:

The record of the applicant is, of course, a starting point, but in every case the final decision is also predicated upon subjective factors such as sincerity, appearance, poise, and the ability to understand and articulate conceptual matters. . . . Only during the interview at the Firm is an evaluation of the intangible factors discussed above possible. Thus, the initial impression of the applicant is of necessity a lasting one. Much depends upon the "chemistry" which occurs between the applicant and the interviewers. It is impossible to exaggerate the importance of this aspect of the interview, for a person with excellent grades may well simply fail to impress that small number of persons who have the ultimate selection responsibility.¹⁶³

The assessment center techniques that are increasingly used to select managerial employees represent some of the most sophisticated subjective evaluation systems available. Yet they too rely on evaluation of factors that lend themselves to the expression of bias and the perpetuation of the status quo — factors such as leadership, ability to get along with others, and identification with organizational goals.¹⁶⁴

To the very limited extent that courts have scrutinized upper level systems, they have found the same kind of irrationality and bias that were revealed in many lower level employment systems. One example is *Chance v. Board of Examiners*,¹⁶⁵ one of the first cases in which a court subjected

¹⁶³ G. COOPER, H. RABB & H. RUBIN, *FAIR EMPLOYMENT LITIGATION* 192 (1975) (quoting defendant's brief in *Kohn v. Royall, Koegel & Wells*, 59 F.R.D. 515 (S.D.N.Y. 1973) (No. 72 Civ. 2705), *appeal dismissed*, 496 F.2d 1094 (2d Cir. 1974)).

¹⁶⁴ Yurko, *supra* note 6, describes the use of similarly vague criteria by academic institutions, *id.* at 476-78, as well as of such status-quo-oriented criteria as "the consistency of the faculty member's educational philosophy with that of the department and the institution, and the compatibility of his personality and work habits with those of other members of the department." *Id.* at 477 (footnotes omitted).

¹⁶⁵ 458 F.2d 1167 (2d Cir. 1972).

an upper level employment system to traditional title VII standards. Another is a sex discrimination case challenging New York Telephone's system for selecting managerial candidates. The defendant's subjective process was supposed to assess candidates' capacities for supervision and leadership. The defendant gave significant credit for military experience, because "[t]he military . . . tend[s] to give their officers rather immediate roles of responsibility."¹⁶⁶ Teaching, by contrast, was not considered a significant supervisory experience. When asked how he knew that experience as a military officer was more valuable than experience as a teacher, the employee in charge replied, "I guess I'm paid to make this type of judgment."¹⁶⁷

One of the few other cases striking down an upper level subjective system involved a performance evaluation system for professional level employees in a Mississippi agency.¹⁶⁸ The agency had only recently been transformed from a dual, segregated system into an integrated one. The court noted that a substantial part of the evaluation guide for supervisors, who were to assess the job performance of their subordinates, related to characteristics that plainly lent themselves to bias — characteristics such as public acceptance, appearance and grooming, personal conduct, outlook on life, and ethical habits.¹⁶⁹ In another case, in which the *Reader's Digest* was charged with sex discrimination, discovery of personnel records revealed that male employees were sometimes given raises when their families expanded, the rationale being need. The records revealed no evidence that female employees received favorable consideration for raises when they had children.¹⁷⁰

Only if courts apply the demanding business necessity test will upper level systems be exposed to critical inspection. The kinds of questions that courts learned to ask in lower level cases can be put usefully to upper level employers. Have they decided what they think the job entails and what constitutes effective performance? What basis do they have for concluding that their systems are selecting the best employees? Have

¹⁶⁶ *Leisner v. New York Tel. Co.*, 358 F. Supp. 359, 365 (S.D.N.Y. 1973) (quoting testimony of defendant's personnel supervisor).

¹⁶⁷ *Id.* at 369 (quoting testimony).

¹⁶⁸ *Wade v. Mississippi Coop. Extension Serv.*, 372 F. Supp. 126 (N.D. Miss. 1974), *aff'd in part, rev'd in part, vacated in part, and remanded*, 528 F.2d 508 (5th Cir. 1976).

¹⁶⁹ *Id.* at 142.

¹⁷⁰ Interview with George Cooper in New York City (Sept. 13, 1980). Mr. Cooper was counsel to plaintiffs in *Smith v. Readers Digest Ass'n*, 15 Fair Empl. Prac. Cas. (BNA) 1606 (S.D.N.Y. 1974).

they considered alternative selection systems that would be less discriminatory in effect? If so, do they have any basis for concluding that such systems would work less well?

When such questions are asked and employers have provided whatever justifications they can, courts will at least have a better understanding of the systems in question. They may conclude in some cases that the justifications are strong and that the systems should be upheld. In other cases they may conclude that the systems cannot be defended as true merit systems. But courts will at least have information enabling them to decide whether the racially exclusionary impact of upper level systems is justified.

When systems are struck down, employers will be free to devise new systems that can be justified as job-related or that will at least have no racially exclusionary impact. This effort may mean expenditures for validation, but testing experts counsel that more resources should be devoted to developing valid selection systems for the most important jobs.¹⁷¹ Alternatively, striking down the system may create pressure for quotas or racially proportionate hiring. But this pressure will be directly related to the weakness of the justifications for existing systems. Courts are inherently conservative creatures. They are not likely to strike down upper level employment systems unless, upon scrutiny, it becomes clear that the case for a system that excludes blacks has simply not been made.

IV. PRACTICAL IMPLICATIONS OF A SINGLE STANDARD: MAJOR ISSUES IN APPLYING TRADITIONAL TITLE VII STANDARDS TO UPPER LEVEL EMPLOYMENT

Differences between upper and lower level jobs and job systems will require some creativity if traditional title VII standards are to be adapted and applied on the upper level. One important difference is numerical: in the entire nation, there are relatively few bank managers or astronomy professors, and few persons seeking such jobs at any given time. By contrast, a single city may employ many thousands of sanitation workers and may hire hundreds each year. The small number of jobs and applicants on the upper level makes it harder to analyze the racial impact of the selection systems used¹⁷² and harder to prove validity by methods that rely on

¹⁷¹ See, e.g., M. DUNNETTE, PERSONNEL SELECTION AND PLACEMENT 8 (1966).

¹⁷² These problems have been exacerbated by recent Supreme Court decisions requiring greater precision in analyzing adverse racial impact than some courts had been demanding. See, e.g., *Hazelwood School Dist. v. United States*, 433 U.S. 299,

statistical comparisons of test performance with job performance.¹⁷³ Selection systems on the upper level are also likely to be multifactored and discretionary, and therefore more difficult to analyze, while lower level systems often rely on a few absolute, objective requirements — attaining a score above the cutoff on a civil service test, passing a physical examination, or possessing a high school diploma.

Given these differences, the courts must develop new methods for assessing both the job-relatedness¹⁷⁴ and the racial impact of upper level selection systems. They must, for example, be willing to look at the collective impact of selection systems for all of an employer's management positions when those systems share significant common elements. Courts must also be willing to consider challenges to industrywide practices for which meaningful statistics are available. Otherwise, jobs held by very few employees under any particular employer will be immune from scrutiny.¹⁷⁵ When upper level employers use education or experience as important factors in hiring, courts must focus on the impact that use of such factors will necessarily have on blacks as a group, if the number of actual applicants involved is too small for specific impact analysis.

Likewise, courts should take advantage of computer technology and statistical methods to analyze complex subjective systems in order to determine what factors are important in decisionmaking. Thus, even if the employer has no clear policies setting the weight for various objective and subjective factors, analysis of company personnel records can give a fairly clear picture of what role each factor has played.¹⁷⁶

The following Sections discuss some of the major issues

309-10 (1977); *International Bhd. of Teamsters v. United States*, 431 U.S. 324, 348-55 (1977).

¹⁷³ See *supra* note 133.

¹⁷⁴ Synthetic validity, for example, enables employers to validate tests for jobs with few incumbents by combining the results of validity studies for a number of job titles. This technique is based on the assumption that all jobs are merely the collection of identifiable job components. Tests for each of the job components can be validated by developing criteria measures for the job component and performing a validity study using all employees in jobs that contain the same job component as the sample. E. McCORMICK & D. ILGEN, *supra* note 135, at 125-28.

¹⁷⁵ Courts must also explore the usefulness of the Uniform Guidelines' provision that statistical significance may not be demanded under certain circumstances. The Guidelines' four-fifths rule provides that federal agencies will regard as evidence of adverse impact the fact that a passing rate on a selection procedure for a minority group is less than four-fifths the passing rate of the highest group. Uniform Guidelines, *supra* note 13, § 1607.4(D). Absence of statistical significance when small numbers are involved is relevant but not necessarily determinative of whether adverse impact will be found.

¹⁷⁶ The opinion in *Vuyanich v. Republic Nat'l Bank*, 505 F. Supp. 224 (N.D.

courts would face in applying traditional title VII standards to upper level employment.

A. Importance of Class Action Treatment

In lower level cases, courts have granted class certification quite liberally¹⁷⁷ because of the group-oriented nature of the alleged wrong and of any relief that might be appropriate. Moreover, courts have recognized that class treatment may be essential if substantive rights are to be vindicated; most cases would not be brought on an individual basis, given such problems as the costs of litigation and legitimate fears of retaliation.¹⁷⁸ These traditional justifications for class action treatment are equally apposite in upper level job cases. Indeed, such treatment may be even more important to the vindication of substantive rights on the upper level.

Given the many antidiscrimination laws and regulations that exist today, overt and blatant forms of discrimination will be rare, especially on higher levels where the decisionmakers are likely to be more sophisticated and legally astute. Hence the socially significant questions are whether upper level employment systems permit the expression of conscious or unconscious bias, and whether systems that result in racial exclusion can be justified as job-related.

Cases tried on an individual basis will rarely resolve these questions. Analysis of the overall operation of an employment system requires broad proof about how the system has affected groups of candidates and employees. Such evidence might

Tex. 1980), demonstrates how multiple regression and other forms of statistical analysis can be used to create a detailed description of a complex employment scheme. For discussion of the use of these techniques in employment litigation, see D. BALDUS & J. COLE, *supra* note 83; Finkelstein, *supra* note 83; Fisher, *Multiple Regression in Legal Proceedings*, 80 COLUM. L. REV. 702 (1980); Greenfield, *From Equal to Equivalent Pay: Salary Discrimination in Academia*, 6 J.L. & EDUC. 41 (1977); Gwartney, Asher, Haworth & Haworth, *supra* note 30; Note, *Beyond the Prima Facie Case in Employment Discrimination Law: Statistical Proof and Rebuttal*, 89 HARV. L. REV. 387 (1975).

¹⁷⁷ See *supra* pp. 963-64.

¹⁷⁸ See, e.g., 7 C. WRIGHT & A. MILLER, *FEDERAL PRACTICE AND PROCEDURE* § 1762, at 602 (1972).

Rutherglen, *supra* note 52, argues that the general judicial presumption in favor of class action certification should not extend to all title VII cases. This argument, however, is based on the theory that the propriety of class action treatment should turn on the merits of the substantive claim. *Id.* at 724-30. Rutherglen considers class action treatment appropriate for disparate impact claims: "Griggs thus established a theory of liability designed for class wide application." *Id.* at 713. The importance of applying the disparate impact doctrine to upper level cases is discussed at *infra* pp. 1004-06.

theoretically be available to an individual plaintiff who charged that an individual employment decision was the unlawful product of a discriminatory system. As a practical matter, however, evidence in an individual case will be far more limited. When the issue is one individual's job treatment rather than the treatment of a larger class, courts will ordinarily not permit the extensive discovery and proof needed to understand a system's operations.

Moreover, cases brought on an individual basis will, under current legal doctrine, almost inevitably be lost even if the courts permit proof about the operation of the overall system. And if the individual case is lost, the court will have no reason to assess the legality of the employment system as a whole.

A look at what typically happens in upper level cases tried on an individual basis reveals why this is so. Most of these cases have been based on a disparate treatment theory designed to prove covert intentional discrimination. Ordinarily the plaintiff in such a case must demonstrate that he was passed over for a white applicant of comparable or lesser qualifications; from this an inference of discriminatory motive can be drawn. The employer can rebut this inference by offering a nondiscriminatory reason for the employment decision; the plaintiff can then prevail only by proving that the reason is a mere "pretext," a cover for intentional discrimination. The entire case depends on there being a systematic, predictable method of employment decisionmaking, so that a discriminatory decision will stand out as aberrational. These requirements might be satisfied in a lower level case: an applicant for a job as a truck driver might well be able to show that an employer hires all candidates satisfying certain minimal requirements; the employer's defense that the applicant was rejected because of inadequate performance with a previous employer might easily be rejected as a pretext if the employer did not ordinarily check applicants' references. By contrast, a lawyer who applied for work with a law firm or a teacher who applied for an assistant professorship would find it hard to prove that an employment rejection was aberrational and hence presumptively discriminatory. The court could not take a candidate's possession of certain qualifications as evidence that the candidate would ordinarily be hired, and the law firm and the university would argue that their decisions reflected overall subjective assessments of many factors — assessments that a court could not easily unpack or challenge.

On the other hand, plaintiffs may be able to prove that the *overall* operation of a particular system reflects intentional bias. Thus, when a bank's black and white employees are

looked at as groups and when productivity-related factors such as age, experience, and education are controlled for, differential treatment of blacks and whites may indicate that the bank's subjective evaluation process is operating in an intentionally discriminatory way. However, even if this kind of evidence is accepted in an individual case, it will be hard to prove bias against particular individuals unless the bias is blatant: courts are reluctant to stand in the place of the employer and to try to make the difficult determination of how the overall subjective assessment process would have come out had there been no improper motive.¹⁷⁹

Even if cases brought on an individual basis could be won, they would rarely be worth the expense. A glance at the 170-page opinion in *Vuyanich v. Republic National Bank*,¹⁸⁰ a challenge to a bank's employment system, and at the evidence presented in the case by statisticians, econometricians, and industrial psychologists gives some idea of the battle of experts that these cases have become. It is little wonder that some courts have expressed concern at the spectacle of trials lasting for weeks, following years of discovery, and involving a multitude of statistical and other experts and seemingly endless testimony about the credentials of a single candidate.¹⁸¹ These cases are just not worth it, from the point of view of litigants, courts, or society generally, unless broad issues related to the legality of entire job systems are to be considered. Class action treatment is designed to deal with such issues.

Moreover, class action treatment enables courts to avoid the particularized intervention in employer decisionmaking required in individual cases. When courts have determined that selection systems have an unjustified disparate impact, they have generally ordered reform without deciding how the reformed system should look. They have rarely done more than outline broad goals — for instance, that the system have no disparate impact or that it be validated; typically, they have

¹⁷⁹ While most upper level cases brought as individual actions have relied on disparate treatment doctrine, similar problems would be encountered in an individual case relying on a disparate impact doctrine. Even if the plaintiff were allowed to prove that the employer's system had an unjustified impact, the plaintiff would have to show that he or she was qualified for the job in order to prevail. Courts would probably be reluctant to dispute the employer's assertion that the plaintiff was not adequately qualified.

¹⁸⁰ 505 F. Supp. 224 (N.D. Tex. 1980).

¹⁸¹ See, e.g., *Lieberman v. Gant*, 474 F. Supp. 848 (D. Conn. 1979), *aff'd*, 630 F.2d 60 (2d Cir. 1980); *Johnson v. University of Pittsburgh*, 435 F. Supp. 1328 (W.D. Pa. 1977).

outlined principles for how the class as a whole should be treated, but have been able to avoid adjudicating particular candidates' qualifications. The details of individual relief have to a great extent been left to negotiations between the parties.

In addition, liberal class action treatment is required because of the small numbers of candidates for many upper level jobs. Proof of racial impact may be found inadequate unless the numbers are large enough to permit statistically significant conclusions. Class action challenges to a broad range of employer practices must be allowed so that the numbers are large enough to expose existing patterns.

An example illustrates the importance of class action treatment in solving the small-numbers problem. A bank might have twenty job titles on the managerial level, with five hundred employees. In a given job title, such as that of credit analyst, there might be two dozen employees. (In analyzing the impact of this selection process, a court must focus on selections after 1965, because it was only then that discrimination became unlawful under title VII.) If the turnover of credit analysts is low, the numbers of employees hired or rejected will be too small to reveal any disparate impact, except in extreme cases. However, if the bank uses a similar system to make hiring, promotion, and salary decisions for its entire managerial staff and if a challenge to the entire employment scheme is allowed, disparate impact — if any — will be revealed far more easily. For example, a salary analysis, controlled for years of education and relevant experience, could be done of all managerial employees. This analysis might reveal that blacks receive lower salaries than comparably qualified whites, a phenomenon that would be invisible in examining any single job title.

As a practical matter, the class action decision will often determine whether the court sees the larger picture. If the court denies class treatment altogether, an individual seeking a job as a credit analyst will likely be permitted to present evidence only about how credit analysts are selected. If the court grants class action treatment but limits the class to those seeking credit analyst positions, proof will probably be similarly limited. Only if the court grants broad class action treatment, allowing a challenge to the bank's entire system of hiring and promoting managers, is the larger picture likely to be considered legally relevant.¹⁸²

¹⁸² The Supreme Court has granted certiorari on the related issue of the propriety of permitting an employee charging promotion discrimination to represent the class of persons allegedly discriminated against in hiring. *General Tel. Co. v. Falcon*, 59

B. Importance of Disparate Impact Analysis as Compared with Disparate Treatment Analysis

It is striking how many upper level job discrimination cases are analyzed as disparate treatment cases and dismissed without discussion of whether a disparate impact case has been made out. Title VII's significance depends on an understanding of the relationship between the two doctrines and on court willingness to apply disparate impact analysis.

Analytically, plaintiff's *prima facie* case under a disparate treatment theory looks much the same as under a disparate impact theory, at least when the challenge is to an employment system rather than an individual decision. In both, the plaintiff must show that the employment system treats blacks differently from comparably qualified whites.¹⁸³ But the weight of the burden the employer must bear once the plaintiff makes out a *prima facie* case depends on which doctrine is applied. Under the disparate treatment doctrine, the employer need only explain the differential treatment, and unless the plaintiff can show that this explanation is a pretext masking intentional discrimination, the plaintiff loses.¹⁸⁴ Under the disparate impact theory, by contrast, the employer must prove the job-relatedness and business necessity of his selection devices.

The difference between these burdens on the employer will often be the difference between winning and losing. Rational explanations abound for why blacks are assigned to less desirable jobs or receive lower salaries. Proof that business necessity demands that they be treated this way is a different matter. The workings of a typical upper level employment case illustrate the relationship between the two doctrines and the central significance of impact analysis. Ordinarily it will be easy for plaintiffs to show that blacks do less well than whites in terms of hiring, job placement, and salary. To make out a traditional disparate treatment case, however, plaintiffs must also demonstrate that blacks are treated differently from whites with *comparable qualifications*. This might be done through use of objective, quantifiable data concerning years of educa-

U.S.L.W. 3385 (U.S. petition for certiorari filed Sept. 21, 1981) (describing questions presented), *cert. granted*, 50 U.S.L.W. 3465 (U.S. Dec. 7, 1981) (No. 81-574).

¹⁸³ See *supra* pp. 964-65.

¹⁸⁴ See *supra* pp. 964-65 & note 58. Under disparate treatment analysis, once the plaintiff has produced evidence implying discriminatory motive, the employer's only obligation is to "articulate" a nondiscriminatory explanation. See *Texas Dep't of Community Affairs v. Burdine*, 450 U.S. 248 (1981); *Board of Trustees v. Sweeney*, 439 U.S. 24, *vacating and remanding per curiam* 569 F.2d 169 (1st Cir. 1978). *Burdine* makes clear that the employer's explanation must take the form of admissible evidence, 450 U.S. at 255, and must be clear and reasonably specific. *Id.* at 258.

tion, years of related experience, and the like. Plaintiffs can then show that, even when such factors are controlled for, blacks are treated less well than whites.

This showing should be sufficient under traditional disparate treatment analysis to place the burden of explanation on the defendant. The defendant may then theorize that discrepancies between the treatment of blacks and whites are explained by the fact that whites are more likely to possess advanced degrees in business-related fields. Or the defendant might theorize that differences in quality of performance account for the differential treatment. The defendant would attempt to show that, if these additional factors are controlled for, the treatment of blacks and whites is equivalent.

The crucial issue is what should happen at this stage of the litigation. Plaintiffs can go on to the next stage of disparate treatment analysis and try to demonstrate that defendant's explanation is a mere pretext concealing intentional discrimination. Such a demonstration is not easy to make, however. If the explanation is not shown to be a pretext, the question is whether the case should be dismissed. It should not, if courts mean to apply the *Griggs* doctrine to upper level employment systems. Defendants' rebuttal to the disparate treatment case has simply fleshed out a previously obscure employment scheme, indicating more specifically which policies are responsible for blacks' doing less well. Under *Griggs*, any such policies are suspect: blacks are presumed capable of doing as well as whites; therefore, it is up to the employer to justify any policies that have a disparate impact. For example, if the defendant's explanation of disparate treatment in the hypothetical above rests in part on the fact that business-related degrees are used as employment criteria and that blacks are less likely to have such degrees, the defendant should have to prove the job-relatedness of such degrees. Similarly, if the defendant's explanation rests on subjective evaluations of employee performance, the bank should have to prove that the evaluations were accurate or that the evaluation system was job-related.¹⁸⁵

¹⁸⁵ See D. BALDUS & J. COLE, *supra* note 83, ¶ 1.2, at 46-47 (under traditional legal standards, defendants should have burden of justifying factors produced in explanation of disparate treatment).

In *James v. Stockham Valves & Fittings Co.*, 559 F.2d 310 (5th Cir. 1977), *cert. denied*, 434 U.S. 1034 (1978), the court followed the approach outlined in text. The defendant had attempted to explain disparate treatment by a regression analysis relying on employees' merit rating and educational level. Merit rating was rejected as a suspect subjective process. Educational level was rejected on the grounds of disparate impact and lack of apparent job-relatedness. *Id.* at 332. Similarly, in a

Failure to apply impact analysis in this manner takes the teeth out of title VII on the upper level; employers will ordinarily have apparently rational explanations for the differential treatment of black employees — explanations resting, for example, on differences in experience and education. The entire point of *Griggs* was to demand that employers produce something more than seemingly reasonable justifications for policies with a negative impact on blacks.

Courts may sometimes fail to apply disparate impact analysis because plaintiffs' initial case rests on a disparate treatment theory. This initial plaintiff strategy may be sensible when the employment scheme at issue is obscure and all that is clear is that it results in differential treatment — as is often the case in upper level employment. Plaintiffs cannot always know, when they first file suit, how defendants will eventually describe their employment policies; it is not possible for plaintiffs to show that unknown policies have a disparate impact. Courts must recognize that one important function of a disparate treatment challenge is to force defendants to articulate their policies so that analysis of their legality is possible.¹⁸⁶

C. Advantages of Griggs Approach in Assessing Legality of Subjective Selection Systems

Courts striking down lower level subjective systems have often suggested that systems based on purely objective criteria would be preferable. It is not clear that this solution is particularly good on the lower level, and it certainly does not

professional level case involving provision of agricultural extension services, the court held that, when the defendant's explanation of disparate treatment was based in part on a performance evaluation system, the system was subject to *Griggs* analysis. Because the system had a disparate impact and had not been validated, the explanation was rejected. *Wade v. Mississippi Coop. Extension Serv.*, 372 F. Supp. 126, 142-43 (N.D. Miss. 1974), *aff'd in part, rev'd in part on other grounds, vacated in part on other grounds, and remanded*, 528 F.2d 508 (5th Cir. 1976).

It should not be enough to offer, as did one defendant, an unsubstantiated theory that women have less motivation for advancement. See *EEOC v. Akron Nat'l Bank & Trust Co.*, 22 Fair Empl. Prac. Cas. (BNA) 1665, 1679 (N.D. Ohio 1980) (rejecting such a defense). Instead, the employer should have to *prove* that women are less qualified for advancement.

¹⁸⁶ Unless the disparate treatment and disparate impact doctrines work together in this fashion, it is hard to make sense of the coexistence of two theories for proving discrimination that place such different burdens of justification on the employer. This problem is noted and an alternative solution presented in Belton, *Burdens of Pleading and Proof in Discrimination Cases: Toward a Theory of Procedural Justice*, 34 VAND. L. REV. 1205, 1266-73 (1981). Professor Belton argues that, in disparate treatment cases, the burden of *persuasion* should shift to the defendant, once plaintiffs have made out a *prima facie* case. However, the Supreme Court ruled to the contrary in *Texas Dep't of Community Affairs v. Burdine*, 450 U.S. 248, 258 (1981).

make much sense on the upper level. Few would argue, for example, that business managers should be promoted solely on the basis of seniority or that academics should be hired on the basis of the number of hours taught or pages published.

Another alternative would be insistence on procedural reform of subjective systems. Many courts have suggested that unlawful systems could be cured by the development of specific guidelines to control the exercise of discretion, by requirements that job openings be advertised, and by addition of new decisionmakers and layers of review.¹⁸⁷

Procedural reform may be an important first step. It may help to control the bias of individual decisionmakers and to open up and expose subjective decisionmaking for review by courts. But this route has significant limitations. First, procedural controls cannot provide complete protection against conscious and unconscious bias. As long as discretionary judgment remains an essential part of the system — as it probably will in upper level selection — there will be room for the expression of bias.¹⁸⁸ There is a danger, therefore, in looking to procedural reform as an adequate substitute for validation.

Moreover, some procedural reforms might be thought to interfere with effective decisionmaking. Thus, courts often refer to the need for guidelines specifying the weight to be given various factors. But predictive judgments about how people will perform complex jobs may best be made by entrusting an intelligent decisionmaker with discretion to make

¹⁸⁷ See, e.g., *Frink v. United States Navy*, 16 Fair Empl. Prac. Cas. (BNA) 67, 70-71 (E.D. Pa. 1977), *aff'd mem.*, 609 F.2d 501 (3d Cir. 1979), *cert. denied*, 445 U.S. 930 (1980); *Miller v. Continental Can Co.*, 13 Fair Empl. Prac. Cas. (BNA) 1585, 1602-03 (S.D. Ga. 1976).

¹⁸⁸ Guidelines specifying the weight to be accorded particular factors have been shown to have limited impact; decisionmakers still tend to give an assessment based on their overall judgment. Cf. E. McCORMICK & D. ILGEN, *supra* note 135, at 77-78 (the "halo effect" causes evaluation of the candidate on many criteria to be influenced by the decisionmaker's evaluation of the candidate on one criterion); Cooper, *Ubiquitous Halo*, 90 PSYCHOLOGICAL BULL. 218 (1981) (discussion of the halo effect and procedures to minimize it). Even when asked to evaluate performance on a very simple task, decisionmakers may be influenced by race. In one laboratory study, black and white job applicants were filmed stacking cans onto shelves; the only difference in their performance was that some stacked more swiftly than others. Blacks who performed as swiftly as whites were found by white raters to be less qualified than their white counterparts. Hamner, Kim, Baird & Bigoness, *Race and Sex as Determinants of Ratings by Potential Employers in a Simulated Work-Sampling Task*, 59 J. APPLIED PSYCHOLOGY 705 (1974). Pooling judgments and adding levels of review limit the impact only of particular biased individuals. If the alleged bias at issue is systemic — if the concern is that white decisionmakers as a group will judge blacks differently than they would judge whites — then adding white decisionmakers will do little good.

an overall assessment. In fact, employers at the upper level have rarely chosen to use elaborate weighting schemes, presumably because no one is sure exactly what factors make for success and how important each is.

One of the great advantages of the *Griggs* approach is that, by focusing on results and proof of job-relatedness, it gives the employer maximum freedom to design employment systems; the sole constraint is that the systems must not exclude blacks unless there is a powerful justification for their doing so. Thus, the employer can attempt to validate its subjective system without having to moderate it in any way.¹⁸⁹ If unable to validate its system, the employer is free to develop another system that is designed to be job-related or that does not exclude blacks.

*D. Assessing Proof of Job-Relatedness and Business
Necessity on the Upper Level*

1. *Limitations of the Testing Experts.* — The testing experts have developed extremely helpful methods for analyzing existing employment systems, but there are limits on their ability to resolve fundamental questions about whether racially exclusionary systems can be justified. In lower level job cases, the role of testing experts has been largely to debunk traditional employment systems. On the upper level, however, experts can be expected to play a significant role both in defending current employment schemes and in creating new ones designed to withstand legal challenge.

It is disturbing that the "new" employment systems designed by testing experts often look similar to the systems earlier struck down by the courts as violative of title VII, and result in similar black failure rates. This result would be defensible if the new test development process were to be trusted.¹⁹⁰

¹⁸⁹ The employer may have to analyze how the system works for purposes of validation, but the operation of the system will not have to be altered.

¹⁹⁰ Courts have found reason to be suspicious of the design of new selection procedures. In *Guardians Ass'n of N.Y. City Police Dep't, Inc. v. Civil Serv. Comm'n*, 484 F. Supp. 785 (S.D.N.Y.), *aff'd in part, rev'd in part on other grounds, and remanded*, 630 F.2d 79 (2d Cir. 1980), *cert. denied*, 101 S. Ct. 3081 (1981), the court struck down a selection test for police officers. It noted that the test had a disparate impact and that the defendant's failure to follow appropriate procedures for test development was part of a long pattern of discriminatory testing. *Id.* at 798. In *United States v. San Diego County*, 21 Fair Empl. Prac. Cas. (BNA) 402 (S.D. Cal. 1979), the court enjoined the use of a new sheriff sergeant's exam designed to replace another exam found unlawful. The court rejected the new exam, in part because it did not test for abilities found to be important in a job analysis. *Id.* at 414.

However, courts should be wary of accepting without question the solutions and defenses proposed by testing experts. First, these experts are likely as a group to see things primarily from the employer's perspective. The experts who designed the validity studies and employment schemes at issue in a title VII lawsuit will ordinarily have been hired by the employer. Plaintiffs will of course have an opportunity to present experts, but their testimony will ordinarily be limited to criticizing studies and employment schemes produced by the employer's experts. Moreover, the testing experts' professional bias makes them likely to be employer oriented. It is the business of industrial psychologists and other testing experts to develop improved selection systems for employers. They inevitably tend to be committed to the rationality of the testing devices they are trained to develop.

A second reason that courts should be reluctant to defer too much to the testing experts is that these experts are, after all, simply technicians; their expertise gives them no basis for making the value choices that determine the legality of selection systems. The danger is that courts will accept the myth of expertise and avoid their own responsibilities to enforce the mandate of title VII. For it is the courts that must decide whether employers, with the aid of their experts, have adequately explained why their job performance goals have to be defined and their selection systems designed such that blacks are significantly excluded from participation.

2. *Problems and Potential of Validation on the Upper Level.* — (a) *Job Analysis.* — Analysis of what the job does or should entail is crucial to any attempt to design job-related selection systems and to any demonstration that a particular system is in fact job related. We must know the kind of performance we are looking for before we can begin rationally to design systems that select people capable of that performance.¹⁹¹

¹⁹¹ The testing literature recognizes the importance of job analysis. See, e.g., L. CRONBACH, *ESSENTIALS OF PSYCHOLOGICAL TESTING* 407 (3d ed. 1970) (job analysis is the first task in performing a criterion-related validation study); DIVISION 14 PRINCIPLES, *supra* note 12, at 4, 7 (a job analysis is essential for proving content validity and, in most cases, criterion-related validity); E. MCCORMICK, *JOB ANALYSIS* 240-71 (1979); see also Uniform Guidelines, *supra* note 13, § 1607.14(A), (B)(2), (B)(3) (validity studies must be based on a review of information about the job, which ordinarily must include a job analysis).

The case law also indicates that job analysis is the critical first step. See, e.g., *United States v. City of Chicago*, 549 F.2d 415, 429 (7th Cir.), cert. denied, 434 U.S. 875 (1977); *Vulcan Soc'y of the N.Y. City Fire Dep't v. Civil Serv. Comm'n*, 490 F.2d 387, 395 (2d Cir. 1973).

A meaningful job analysis will be difficult for many upper level jobs, because we are uncertain what constitutes good performance. But without an adequate job analysis, validation is impossible. If we concede, for example, that we have very little idea what it means to be a good lawyer and that we consequently cannot analyze how important to good lawyering are the analytical skills measured by law school examinations, there can be no way of validating a law firm's reliance on law school grades for selecting lawyers. Similarly, if we concede that we do not know what makes for effective teaching, as did a lower court in a case upholding South Carolina's use of the National Teacher Examination (NTE), there can be no way of measuring the validity of that examination as a device for selecting teachers.¹⁹²

(i) *Inadequacy of Job Description Approach.* — The task of job analysis should not be seen as simply one of job description. The purpose of the job selection enterprise is to select those most qualified to do the *best* job, not simply to perpetuate current levels of performance. The purpose of validation in the context of a racial discrimination claim is to assess whether particular selection systems are necessary to promote good performance despite their negative impact on blacks. The latter purpose is not served if, for purposes of validation, one describes jobs simply by describing what present job incumbents do. An appropriate analysis of the job of police officer in a department whose officers engage in brutality would not consist simply of a description of the way current officers perform their jobs.

The job description approach is likely to lead to the creation of "new" selection systems that perpetuate the racial exclusion characterizing the old system. Thus, if almost all the incumbents of a job are whites with middle class backgrounds, a job analysis that focuses solely on the way they perform their jobs is likely to reflect characteristics common to middle class whites. Selection devices designed to test for those characteristics are likely to continue to select whites at disproportionate rates.

¹⁹² *United States v. South Carolina*, 445 F. Supp. 1094, 1108 n.13 (D.S.C. 1977), *aff'd mem. sub nom. National Educ. Ass'n v. South Carolina*, 434 U.S. 1026 (1978). Without a job analysis, there is no sense of what combination of skills makes for effective job performance overall. Even if one assumes that the skills measured by law school examinations or by the NTE are important to some part of effective job performance, one does not know how large that part is, what other skills may be important, or whether ability to perform on the examinations correlates with ability to perform in ways that the examinations do not purport to measure.

Although the importance of not limiting job analysis to mere job description seems obvious, the point has received little attention. Indeed, the task of job analysis has been routinely described in the testing literature as one simply involving description.¹⁹³ When defendants have presented evidence of validation efforts, they have usually relied primarily on a review of what current jobholders are doing.¹⁹⁴ The "new" selection systems spawned by this type of job analysis have often had the same kind of adverse impact on blacks as the systems previously struck down.¹⁹⁵ And of course, this result tends simply to confirm the views of many that blacks are less capable of performing well on the job. But it should be clear that, if job analysis relies on the activities and views of job incumbents, a selection system based on that analysis may well select a candidate group similar to the incumbents, with a similar racial composition.

This problem surfaced at one stage of the *Chance* case,¹⁹⁶ a suit challenging the criteria for selecting school principals and other supervisors in New York City's public school system. Plaintiffs charged that the selection system had an adverse racial impact and was not job related; they alleged that the system failed to select the most qualified candidates because it ignored such characteristics as a principal's ability to relate to the local community. After the old system was struck down, defendants hired experts to conduct job analyses in an attempt to develop a legally defensible selection system. The analyses consisted largely of descriptions of what the incumbent school supervisors — almost all whites — did on the job and of which tasks they considered most important. Plaintiffs successfully opposed court approval of a plan to develop ex-

¹⁹³ See, e.g., E. McCORMICK, *supra* note 191, at 48; E. McCORMICK & D. ILGEN, *supra* note 135, at 37-46; McCormick, *Job and Task Analysis*, in HANDBOOK, *supra* note 101, at 651, 652-53.

¹⁹⁴ The most common methodology has been to poll current jobholders and their supervisors about which tasks they feel are the most important. See, e.g., Guardians Ass'n of the N.Y. City Police Dep't, Inc. v. Civil Serv. Comm'n, 484 F. Supp. 785 (S.D.N.Y.) *aff'd in part, vacated in part, and remanded*, 630 F.2d 79 (2d Cir. 1980), *cert. denied*, 101 S. Ct. 3083 (1981); *Chance v. Board of Examiners*, 330 F. Supp. 203, 216 (S.D.N.Y. 1971), *aff'd*, 458 F.2d 1167 (2d Cir. 1972).

¹⁹⁵ See, e.g., Firefighters Inst. for Racial Equality v. City of St. Louis, 549 F.2d 506 (8th Cir. 1977), *aff'g in part, rev'g in part, and remanding in part* United States v. City of St. Louis, 410 F. Supp. 948, 953 (E.D. Mo. 1976), *cert. denied*, 434 U.S. 819 (1977); Guardians Ass'n of the N.Y. City Police Dep't, Inc. v. Civil Serv. Comm'n, 484 F. Supp. 785 (S.D.N.Y.), *aff'd in part, vacated in part, and remanded*, 630 F.2d 79 (2d Cir. 1980), *cert. denied*, 101 S. Ct. 3083 (1981); Arnold v. Ballard, 390 F. Supp. 723, 731-33 (N.D. Ohio 1975).

¹⁹⁶ *Chance v. Board of Examiners*, 458 F.2d 1167 (2d Cir. 1972).

aminations based on these analyses;¹⁹⁷ they argued that a job analysis that looked primarily to the activities and value judgments of white incumbents would be hopelessly flawed.

Unless the courts insist otherwise, job analysis may well continue to be seen as mere job description, because employers and their experts will be naturally biased in favor of keeping the job as it has been. The testing experts' only real claim to expertise in this area is in techniques of describing what jobs entail. They certainly are not experts in deciding what a school principal or lawyer or business manager *ought* to be doing. Unless instructed otherwise, the testing expert would probably leave these value choices to the employer and would accept the status quo. And the employer is likely to have a bias in favor of the way it has been defining performance. Even an employer attempting in good faith to test its previous assumptions is likely to think that a descriptive job analysis makes perfect sense if the "experts" treat it as the norm.

Courts are likely to resist the notion that they should interfere with the employer's determination of how job performance goals should be defined. They are likely to protest that they are not experts in deciding what the ideal lawyer, manager, or academic should look like.¹⁹⁸ But it is to the courts that our legal system has entrusted responsibility for weighing employers' justifications for running their businesses as they traditionally have against the values embodied in title VII's prohibition of racial discrimination. Courts cannot fully enforce that prohibition if they refuse to consider whether employers are justified in defining jobs of key importance to both blacks and whites by looking only to white views of what

¹⁹⁷ Transcript of hearing and of decision from the bench, at 13, 44, 45, 56, 64, *Chance v. Board of Examiners*, No. 70 Civ. 4141 (S.D.N.Y. June 1, 1976).

At a later stage of the case, defendants proceeded to develop "new" examinations based on job analyses prepared by similar methods. Their racial impact was so harsh that the chancellor of the city school system initially refused to promulgate the lists of successful candidates. *Testing Principals and Principles*, N.Y. Times, Aug. 11, 1981, at A16, col. 1 (editorial). A suit is now pending challenging the new examination system on grounds similar to those on which the original suit was based. *Marchiarola v. Board of Supervisors*, No. 81 Civ. 4798 (S.D.N.Y. amended complaint filed Sept. 24, 1981).

¹⁹⁸ See *Lynn v. Regents of Univ. of Cal.*, 21 Empl. Prac. Dec. (CCH) ¶ 30,558 (C.D. Cal. 1979) (court declines to intervene despite evidence that negative tenure decision reflected disapproval of feminist nature of candidate's work); *Scott v. University of Del.*, 455 F. Supp. 1102, 1126 (D. Del. 1978) (challenge to doctoral requirement for academic promotion held inappropriate because "the University's choice of mission is not a subject for judicial review"), *aff'd in part and vacated in part on other grounds*, 601 F.2d 76 (3d Cir.), *cert. denied*, 444 U.S. 931 (1979).

those jobs should be. This does not mean that courts should step in to define job performance goals in the first instance. Employers would continue to play that role. But it does mean that, in reviewing the legality of an employment system that has a racially exclusionary effect, courts should consider whether the employer's job definition is partly responsible for that effect and, if so, whether the definition is defensible.

Courts have occasionally reviewed job definitions in this manner. In *Diaz v. Pan American World Airways, Inc.*,¹⁹⁹ the employer's defense of its females-only flight attendant policy relied on proof that women were for various psychological reasons better suited than men to perform a function that the employer argued was central to the flight attendant job — to make passengers feel comfortable. The court struck down the defense, finding that this function was simply not a part of the court's own concept of a flight attendant's main responsibility — "to transport passengers safely."²⁰⁰ The fact that women apparently performed the flight attendant job as defined by the employer better than men was deemed irrelevant, because the court decided that the nonsafety aspects of the job were "tangential to the essence of the business."²⁰¹ In effect, the court overrode the employer's job definition, because that definition excluded men from the job for reasons that did not seem particularly weighty. Obviously the court was no expert on the flight attendant's job, nor would it ordinarily be appropriate for a court to decide how that job should be defined. But it is the court's function to weigh the cost of excluding protected groups against the cost of interfering with employer choices;²⁰² for that reason, the court could properly say that the essence of a flight attendant's job was or should be related to passenger safety.

(ii) *Race as Job-Related.* — Those engaged in designing job analyses should not only consider whether the job definition will perpetuate the exclusion of blacks; they should also be willing to consider blackness in a positive sense. They should determine whether the job is one in which the inclusion of racial minorities is important to improving the quality of performance. Many have argued, for example, that there is

¹⁹⁹ 442 F.2d 385 (5th Cir.), *cert. denied*, 404 U.S. 950 (1971).

²⁰⁰ *Id.* at 388.

²⁰¹ *Id.*

²⁰² Courts have traditionally engaged in this kind of balancing in interpreting title VII. The business necessity test, for example, requires courts to decide whether employment policies represent interests that are important enough to outweigh their racial impact. *See, e.g., Robinson v. Lorillard Corp.*, 444 F.2d 791, 799 (4th Cir.), *cert. dismissed*, 404 U.S. 1006 (1971).

a need for more black lawyers who are interested in working with minority communities, for more black teachers who are capable of effectively communicating with black students and their parents, and for more black academics who are interested in focusing their teaching and scholarship on problems particularly significant to minority groups.²⁰³

The issue whether blackness may be considered job-related in a positive sense under certain circumstances has rarely been the focus of employment discrimination litigation. Race-conscious employment plans tend to be analyzed in terms of whether they can be justified in the name of affirmative action, as compensation for previous discrimination. Courts need to consider in addition whether such plans can be justified because performance in certain jobs in today's society will be enhanced by increasing the percentage of blacks holding those jobs.

Court decisions occasionally indicate that race can legitimately be considered job-related. In *Regents of the University of California v. Bakke*, Justice Powell recognized that under some circumstances the government interest in providing health care to disadvantaged groups would justify a race-conscious admissions program.²⁰⁴ Similarly, in a district court case involving a police force, the court argued as follows:

[A]ll citizens profit when the city achieves a racially integrated police force of qualified individuals who are knowledgeable of the diverse problems of different ethnic groups and who are not prey to destructive hostility from minorities who feel excluded from full participation in city government life. Clearly, the general harmony of the community is enhanced by the city's obtaining a police force representative of its population.²⁰⁵

Other courts have given similarly explicit recognition to this concept in relation to the hiring of police and firefighters.²⁰⁶

²⁰³ The concept of race as a "merit-related" qualification is discussed in *R. FULLINWIDER, THE REVERSE DISCRIMINATION CONTROVERSY* 78-83, 86-88 (1980); Fallon, *supra* note 24, at 819, 842-43, 860-62 (1980).

²⁰⁴ *Regents of the Univ. of Cal. v. Bakke*, 438 U.S. 265, 310-11 (1978) (opinion of Powell, J.).

²⁰⁵ *Officers for Justice v. Civil Serv. Comm'n*, 371 F. Supp. 1328, 1330-31 (N.D. Cal. 1973).

²⁰⁶ See, e.g., *Detroit Police Officers' Ass'n v. Young*, 608 F.2d 671, 696 (6th Cir. 1979), *cert. denied*, 101 S. Ct. 3079 (1981); *League of United Latin Am. Citizens v. City of Santa Ana*, 410 F. Supp. 873, 896-97 (C.D. Cal. 1976); Note, *Race as an Employment Qualification to Meet Police Department Operational Needs*, 54 N.Y.U. L. REV. 413 (1979); see also *Talbert v. City of Richmond*, 648 F.2d 925 (4th Cir. 1981) (relying on *Detroit Police Officers' Ass'n v. Young* to legitimate goal of racial

In other cases, the way of defining the geographic pools suggests that some notion that race is job related is lurking in the background. The size of the relevant pool is ordinarily calculated in title VII litigation by reference to the area from which the defendants' employees are recruited. In many police and fire cases, however, the courts have held that the appropriate area is instead the jurisdiction served.²⁰⁷ For example, if the population of a city is forty percent black, the court would find suspect any examination process that resulted in a police force less than forty percent black. This reliance on city population statistics when, as is usually the case, officers are recruited from a larger and less black area, makes no sense if the purpose is to assess the impact of the selection devices on candidates. These courts seem to be expressing an unspoken assumption that race is job related and that police and firefighters should reflect the racial composition of the population they serve.

In upholding a prison guard hiring and promotion plan, the California Court of Appeals recognized as legitimate the job-relatedness of race. The plan was designed to create a guard force with a minority percentage at least seventy percent that of the prison population. After noting evidence showing that the plan "would serve inmate-related goals of the Department by improving relationships with prisoners and reducing severe racial conflict and violence within the state prison system," the court upheld the plan.²⁰⁸

There are questions, not dealt with here, about the legality

diversity in police department's upper ranks), *cert. denied*, 50 U.S.L.W. 3547 (U.S. Jan. 11, 1982); *Bridgeport Guardians, Inc. v. Members of Bridgeport Civil Serv. Comm'n*, 482 F.2d 1333, 1341 (2d Cir. 1973) ("[T]he visibility of the Black patrolman in the community is a decided advantage for all segments of the public at a time when racial divisiveness is plaguing law enforcement."), *cert. denied*, 421 U.S. 991 (1975).

²⁰⁷ See, e.g., *Detroit Police Officers' Ass'n v. Young*, 608 F.2d at 688 (6th Cir. 1975), *vacated and remanded*, 431 U.S. 951 (1977); *Boston Chapter, NAACP, Inc. v. Beecher*, 504 F.2d 1017, 1020 n.4 (1st Cir. 1974) (firefighters), *cert. denied*, 421 U.S. 910 (1975); *Afro Am. Patrolmens League v. Duck*, 503 F.2d 294, 299 (6th Cir. 1974) (police); *Erie Human Relations Comm'n v. Tullio*, 493 F.2d 371, 374-75 (3d Cir. 1974) (police); *Vulcan Soc'y of the N.Y. City Fire Dep't v. Civil Serv. Comm'n*, 490 F.2d 387, 398 (2d Cir. 1973); *League of United Latin Am. Citizens v. City of Santa Ana*, 410 F. Supp. 873, 889-90 & n.13, 896-97 (C.D. Cal. 1976) (police and firefighters).

²⁰⁸ *Minnick v. Department of Corrections*, 95 Cal. App. 3d 506, 157 Cal. Rptr. 260 (1979), *cert. dismissed*, 101 S. Ct. 2211 (1981); see also *Porcelli v. Titus*, 302 F. Supp. 726, 732-33 (D.N.J. 1969) (recognizing need for black authority figures with whom pupils could identify and upholding system using race as a factor in promoting teachers to administrative positions), *aff'd*, 431 F.2d 1254 (3d Cir. 1970), *cert. denied*, 402 U.S. 944 (1971).

under title VII and the Constitution of giving positive consideration to blackness in the definition of certain jobs.²⁰⁹ It is important, however, for courts to realize that the issues cannot be reduced to those usually considered in the context of "affirmative action" plans — namely, whether racial preferences can be justified as a means to make up for past injustices. If race is job related in certain instances, this fact provides a different kind of justification for race-conscious employment schemes.

(b) *Validation Techniques*. — Validation of a test involves a demonstration that the test does what it purports to do. Validation of an employee selection system will normally consist of proof that the system successfully predicts performance on the job. The following discussion examines the commonly accepted methods of validation, focusing on some of the major issues likely to arise in validating upper level selection systems.

(i) *Content Validity*. — A test is said to be content-valid with respect to a job when it measures performance of tasks that constitute a relatively complete sample of those called for on the job.²¹⁰ Content validity has generally been considered appropriate for jobs that consist primarily of a few simple tasks. The example commonly given is the typist. If the job consists almost entirely of typing, a test that measures the candidate's typing ability will ordinarily be considered content-valid.

This theory provides little justification for upper level selection systems. Performance in upper level jobs ordinarily involves a number of sophisticated functions, and it is not possible to include an adequate sample of all of them in a test. Indeed, very few of the selection devices used on the upper level purport to provide a sample of the actual job. Interviews, assessments of educational accomplishments and of work performance, and traditional mental aptitude tests, for example, are all devices designed to predict future behavior, based on past or present behavior of a different nature. Assessment center techniques often try to simulate actual job situations. Candidates may be given an in-basket and asked to make decisions about what should be done with each

²⁰⁹ See *Knight v. Nassau County Civil Serv. Comm'n*, 25 Fair Empl. Prac. Cas. (BNA) 1448, 1451 (2d Cir. 1981) (violation of title VII and of Constitution to use black employee's race as a factor in assigning him to minority recruitment). A discussion of related issues is contained in Note, *supra* note 206.

²¹⁰ See, e.g., Uniform Guidelines, *supra* note 13, § 1607.14(C)(1); E. McCORMICK & D. ILGEN, *supra* note 135, at 111; Guion, *Scoring of Content Domain Samples: The Problem of Fairness*, 63 J. APPLIED PSYCHOLOGY 499, 501-02 (1978).

item — drafting an answer to a letter, delegating a task to a subordinate, and the like. Or candidates may be placed in discussion groups and given roles designed to demonstrate their leadership abilities. The goal, however, will usually not be to reproduce in the assessment center a representative sample of the job, but rather to assess a set of abilities that seem likely to be related to future ability to perform. Indeed, there is probably no way that the tasks and real-world pressures involved in any complex job could be replicated, short of giving the candidate a probationary period in the actual job.

The Uniform Guidelines provide an alternative basis for demonstrating content validity. Selection procedures that purport to measure “knowledge, skills, or abilities” — even if they are not representative samples — may be justified by content validity if the knowledge, skill, or ability “is a necessary prerequisite to successful job performance.”²¹¹

Even if one of these theories of content validity seems applicable, numerous problems in demonstrating validity remain. First, any claim for content validity depends on an adequate job analysis.²¹² Second, should employers design tests that are content-valid for limited aspects of an upper level job — say, for minimum required knowledge — that alone would not validate the overall selection system.²¹³ Third, a test is content-valid only if performance on the test is measured fairly and accurately.²¹⁴ Upper level job selection devices are likely to include subjective elements; the measurement problems likely to arise are therefore those typical of all subjective assessment processes.²¹⁵ For example, after observing a candi-

²¹¹ Uniform Guidelines, *supra* note 13, § 1607.14(C)(1). By contrast, the Uniform Guidelines find a content validity strategy inappropriate for selection procedures that purport to measure “traits or constructs,” such as intelligence, aptitude, personality, commonsense, judgment, and leadership. *Id.* The APA has come to similar conclusions. DIVISION 14 PRINCIPLES, *supra* note 12, at 13.

²¹² Uniform Guidelines, *supra* note 13, § 1607.14(C)(2); DIVISION 14 PRINCIPLES, *supra* note 12, at 13.

²¹³ See, e.g., *Firefighters Inst. for Racial Equality v. City of St. Louis*, 549 F.2d 506, 511–12 (8th Cir.) (fire captain examination not content-valid when it tested job knowledge but not supervisory ability, admittedly a key aspect of the job), *cert. denied*, 434 U.S. 819 (1977); *United States v. San Diego County*, 21 Fair Empl. Prac. Cas. 402 (BNA) (S.D. Cal. 1979) (striking down selection test that measured only two of four attributes found necessary to job).

²¹⁴ See, e.g., Guion, “Content Validity” in *Moderation*, 31 PERSONNEL PSYCHOLOGY 205, 208–11 (1978). For example, a typing test would not be a fair selection device, even if it represented a fair sample of the job at issue, if the cutoff score were set so high that it excluded typists who type fast enough to satisfy job requirements.

²¹⁵ See generally Guion, *supra* note 210, at 504–06 (serious problems of potential bias in scoring of content-structured tests).

date's performance, an assessment staff makes subjective judgements about such qualities as leadership ability, initiative, and ability to get along with others.²¹⁶ Even if the performance evaluated is comparable to the performance required by the job, the staff's evaluation process is subjective and must be analyzed like any other subjective test. If it has a disparate impact, it is suspect for the same reasons pointed to in *Rowe* and related cases.²¹⁷ The assessment staff, drawn as it usually is from the supervisory ranks at or above the level to which the candidates aspire, is likely to be disproportionately white compared with the candidates. Even if the staff consists of outside professionals, they take their goals from the employer; their criteria and any formulae they use to weight those criteria will reflect the goals and values of the enterprise and its top management.²¹⁸ And beyond considerations of conscious or unconscious bias, there remains a real danger that this kind of assessment scheme will select the candidates who most resemble the upper level employees in the enterprise; in that case, an existing pattern of racial exclusion may well be perpetuated, and the employer must then demonstrate that the job *requires* this exclusion.²¹⁹

(ii) *Empirical Validity*. — Empirical validation requires an analysis of the relationship between performance on a test or other "predictor" and performance on the job being tested for. Number of years of experience, medical school grades, and IQ scores are all examples of potential predictors. The goal of the validation process is to demonstrate that a purported predictor actually does predict employee performance.

Empirical validation requires a determination of what good performance is and how to measure it. The testing experts refer to this process as the development of appropriate criteria and criteria measures. A validity study then compares employees' performance on a given predictor against their performance on the appropriate criteria measures. If the correlations are high enough to be practically useful — if it seems more

²¹⁶ See, e.g., Hinrichs, *supra* note 144, at 598; Warbois, *Validation of Externally Developed Assessment Procedures for Identification of Supervisory Potential*, 28 PERSONNEL PSYCHOLOGY 77, 79-80 (1975).

²¹⁷ See *supra* p. 974.

²¹⁸ Criteria used in some assessment centers include "company orientation," defined as "identifying the organization's goals and values as reflecting one's own"; and "functional ability," defined as "existing successfully in one's environment." See, e.g., Warbois, *supra* note 216, at 79-80.

²¹⁹ See generally Fallon, *supra* note 24, at 854-56 & n.143 (discussing problems of using content validation strategy when complex jobs are at issue).

likely than not that the predictor is accurately predicting job performance — the predictor is said to be valid.

Criterion-related validation has been the main form of empirical validation to date, and the main alternative to content validation. A third generally recognized technique is construct validation, which focuses on the measurement of certain mental and other capacities — “constructs” — deemed important to future performance. Construct validation is a relatively new concept, and experts disagree about what it means or how it could be shown. However, there seems to be general agreement that construct validity requires empirical validation — that is, demonstration of a relationship between the construct and job performance.²²⁰

Developing appropriate criteria measures for upper level jobs will be difficult. The criteria measures themselves are yet another kind of testing device; they must fairly and accurately measure employee performance if the validity study is to prove anything. One alternative would be to ask employees to take some sort of objective test designed to measure how well they know their job. Or one could use other objective criteria; for a business manager, they might include department productivity, employee absenteeism, or amount of business attracted to the company. If blacks do worse on such objective tests or criteria, the inquiry cannot simply stop there; it must be determined that the criteria are in fact measuring job performance accurately.

Alternatively, some sort of subjective evaluation might be used as a criterion measure. Thus, the employer might try to assess job performance by using performance ratings. Or the employer might look to salary increases or promotions as mea-

²²⁰ Uniform Guidelines, *supra* note 13, § 1607.14(D)(3) (proof of construct validity requires that the relationship between construct and performance be “supported by empirical evidence from one or more criterion-related studies”); Barrett, *Is the Test Content-Valid: Or, Does It Really Measure a Construct?*, 6 EMPLOYEE REL. L.J. 459, 464 (1981) (construct validation requires “empirical evidence showing the relation between the constructs and the work behavior”); *see also* E. MCCORMICK & D. ILGEN, *supra* note 135, at 124 (proof of construct validity requires empirical demonstration that the test accurately measures the construct as well as a complete job analysis showing the role of the construct in the job); *cf.* A. ANASTASI, *PSYCHOLOGICAL TESTING* 151-61 (4th ed. 1976) (construct validity is a broad concept, encompassing the notion of empirical validity). In *Douglas v. Hampton*, 512 F.2d 976, 986-87 (D.C. Cir. 1975), the court rejected one of the few attempts to justify a test by means of construct validity. The court defined construct validity narrowly to require only a showing that a test accurately measures certain constructs. But the court then said that, “in determining whether a showing of construct validity satisfies *Griggs*, the court must also determine whether the constructs are themselves related to job performance.” *Id.* at 986. The court also held that construct validity may be considered only when empirical validity is infeasible. *Id.*

asures of an employee's job performance, when those decisions have been based on subjective processes. Although such subjective measures may appear far more sensible than a purely objective approach for most upper level jobs, the use of subjective measures poses serious problems of fairness and accuracy. Those who subjectively assess an employee's job performance will ordinarily be higher level employees of the same employer. A subjective criterion measure for evaluating job performance may strongly resemble the selection procedures that kept many minorities from being hired in the first place. Indeed, courts analyzing the use of subjective criterion measures for purposes of empirical validation have recognized how suspect they are.²²¹

The testing literature reveals professional concern with the problems criteria measures pose. Experts note that appropriate criteria measures have generally not been developed. They condemn the tendency simply to seize on easily available measures such as promotions and salary raises without regard to their relationship to job performance.²²² One commentator observed: "If the problem of investigating possible predictor bias is difficult, the problem of criterion bias is appalling."²²³

²²¹ *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975). One court condemning validity studies in a case involving lower level craft jobs reasoned as follows: "Although somewhat constrained, the rating process here was still principally subjective. We have in the past condemned the use of such nebulous standards. Where, as here, subjective evaluations are used in the very process of test validation, a similar potential for abuse exists." *Rogers v. International Paper Co.*, 510 F.2d 1340, 1350 (8th Cir.) (citations omitted), *vacated on other grounds and remanded*, 423 U.S. 809 (1975); see also *United States v. City of Chicago*, 549 F.2d 415, 433 & n.24 (7th Cir.) (supervisors' efficiency ratings rejected because procedures for obtaining ratings not sufficient to ensure that they were a good measure of job performance), *cert. denied*, 434 U.S. 875 (1977); *Watkins v. Scott Paper Co.*, 530 F.2d 1159, 1188-90 (5th Cir.) (ruling inadequate a criterion-related validity study because raters were instructed only to judge which employee performed better, a standard subject to supervisory bias), *cert. denied*, 429 U.S. 861 (1976); *League of United Latin Am. Citizens v. City of Santa Ana*, 410 F. Supp. 873, 905-06 (C.D. Cal. 1976) (validation study using performance ratings as criterion measures rejected, in part because of strong evidence of racial bias against Mexican-Americans).

²²² See, e.g., *MANAGERIAL BEHAVIOR*, *supra* note 27, at 101-26, 473-74; R. GUION, *supra* note 27, at 90-91, 117-18; cf. Ash & Kroeker, *supra* note 132, at 485 ("[T]he criterion remains the weak link in the chain [of test validation]."). See generally Dunnette & Borman, *supra* note 143, at 486 (noting extensive writing about the "criterion problem"); Guion, *supra* note 214, at 205-06 (referring to the "tired old discussions of the 'criterion problem'").

²²³ Guion, *supra* note 135, at 815; see also M. MINER & J. MINER, *supra* note 142, at 96 (recognizing potential for bias in using performance ratings as criterion measures); Dunnette & Borman, *supra* note 143, at 489-90 (noting the need for much work in developing bias-free performance-rating systems); Hamner, Kim, Baird &

He argues that it is "impossible to tell whether an observed ethnic difference is evidence of bias or simply descriptive."²²⁴

Courts have a vital role to play in scrutinizing the adequacy of criteria measures. Employers who are persuaded that their selection schemes are sound may want to prove that hypothesis by the least expensive means possible. Given the difficulty of developing criterion measures for upper level jobs, even the good faith employer will be tempted to seize on some existing measure, whatever its problems, and hope that some apparently persuasive correlation can be produced.

Nor can the courts assume that the experts will develop appropriate criteria measures. The testing literature contains reports of numerous validity studies in which correlations were calculated and validity claimed on the basis of measures that seem obviously suspect, with little or no discussion of the problems involved.²²⁵ A 1977 review of the validity studies of assessment centers illustrates some of the problems involved.²²⁶ The authors question the value of these studies, which consistently purported to demonstrate validity. They note the tendency to use criteria such as salary or promotions, which "may have less to do with managerial effectiveness than with managerial adaptation and survival."²²⁷ The authors point out that this approach to criteria measures tends simply to perpetuate the status quo:

"If we are, indeed, embarked on a venture which will lead us to pick people who can get good ratings, especially from big shots, what are the implications for progress in business and societal endeavor? How to succeed in business by satisfying the guardians of the status quo?" . . . If we wish to move away from the status quo, then, we need criteria other than advancement.²²⁸

Bigoness, *supra* note 188, at 709-10 (criterion measures may be infected with the same bias as the predictor).

²²⁴ Guion, *supra* note 210, at 816.

²²⁵ See generally M. MINER & J. MINER, *supra* note 142, at 94, 96 (performance ratings are typically used to arrive at criterion measures, despite recognized potential for bias in use of such ratings).

²²⁶ Klimoski & Strickland, *supra* note 144.

²²⁷ *Id.* at 355.

²²⁸ *Id.* at 358 (quoting Wallace, *How High the Validity?*, 27 PERSONNEL PSYCHOLOGY 397, 404 (1974)). The authors discuss the possibility that the assessment center operates simply as a "policy-capturing" device — as a means of describing how and why people are advanced in a particular organization. *Id.* at 358-60; see also MANAGERIAL BEHAVIOR, *supra* note 27, at 101-11, 125-26 (commenting on the status quo orientation of the criterion measures typically used in managerial validity studies); Hinrichs, *supra* note 144, at 600 (much research is needed to develop accurate measure of performance, since assessment centers are generally predictive of achievement).

The danger is that courts will accept the quick-and-dirty validation studies that seem to have become popular. One common response to the civil rights attack on the use of selection devices that exclude blacks has consisted simply of efforts to demonstrate validity, rather than serious attempts to evaluate their usefulness. Because available measures of job success often bear a remarkable resemblance to the challenged predictor, positive correlations between test performance and performance on the criterion measure may well be demonstrated and validity claimed. A number of courts have been persuaded that this kind of demonstration proves something. The Supreme Court in *Washington v. Davis*²²⁹ found significant a relationship between scores on the entry level test for policemen and scores on examinations given during a training program for the same policemen. In *United States v. South Carolina*,²³⁰ the district court found persuasive a correlation between scores on the challenged National Teacher Examination (NTE) and scores in a teacher training program that the candidates had been engaged in prior to taking the NTE; the Supreme Court affirmed summarily. In neither case was there reason to think that the training program test scores bore any relationship to job performance. In both cases, all that was shown was that people who did well on one test did well on another very similar test.

Accepting such evidence as proof of validity is far worse than simply telling upper level employers that they will be subject to a looser standard. The message conveyed is not simply that blacks are to be excluded, but that they are to be excluded because they have been proved less capable of doing good work. To the extent the proof is phony, the message is both harmful and unfair.

Proper validation may seem too demanding a process. If every measure on which blacks do poorly must be justified, whether it be a predictor or criterion measure, where does the process stop? The point, however, is not that courts should demand the impossible in the way of demonstrably valid criteria measures or that they should strike down as unlawful all employment schemes that cannot be proved valid according to the strictest technical standards. Rather, courts should demand validation and should look hard at studies purporting to prove validity. In the process, they will learn much about why blacks do poorly in particular selection schemes and about

²²⁹ 426 U.S. 229 (1976).

²³⁰ 445 F. Supp. 1094 (D.S.C. 1977), *aff'd mem. sub nom.* National Educ. Ass'n v. South Carolina, 434 U.S. 1026 (1978).

the strengths and weaknesses of employer justifications for those schemes. In lower level job cases, such investigation has often led courts to the conclusion that the employment schemes at issue cannot be justified. It is unlikely that courts have been fooled by the experts' jargon. It is likely instead that investigation has shown the courts the real irrationalities of the employment systems at issue and the extent to which they simply reflect the values and biases of those already inside the system.

3. *Duty to Explore and Adopt Less Discriminatory Alternatives of Equivalent Validity.* — Proof of the validity of an employment scheme that has an adverse racial impact does not end the inquiry. Traditional title VII standards demand that the employer opt for any available alternative system that has a lesser impact, as long as it also serves the employer's job needs.²³¹ Validity simply means that a given selection device is a better predictor of successful performance than a random selection process: there may be several valid systems for predicting performance in any given job, and any one of them is likely to measure only certain of the characteristics that determine job success. The problem is that the characteristics measured by one test may be highly correlated with race, while those measured by another may not.²³²

Recognition of the less-discriminatory-alternative doctrine

²³¹ See, e.g., *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975); *Robinson v. Lorillard Corp.*, 444 F.2d 791, 798 n.7 (4th Cir.), cert. dismissed, 404 U.S. 1006 (1971); *Crockett v. Green*, 388 F. Supp. 912, 920 (E.D. Wis. 1975), *aff'd*, 534 F.2d 715 (7th Cir. 1976). The Uniform Guidelines provide that, if the employer is shown an alternative selection method that appears to have substantial validity and less adverse impact, he must investigate its potential. If two methods have substantially equal validity but one has a less discriminatory result, the employer must choose the latter. Uniform Guidelines, *supra* note 13, § 1607.3(B).

²³² The issues involved in an exploration of less discriminatory alternatives are different from those involved in the "differential validity" and "fairness" inquiries that have been the subject of much debate during the past decade. Differential validity exists when the validity coefficient for one group differs from that for another group. Thus, a test might be a good predictor of performance for whites, but a poor predictor of performance for blacks. Differential validity was virtually dismissed as nonexistent in the early 1970's, but the controversy over its existence has been reopened in recent years.

Fairness inquiries ask a different question: Does a test systematically overpredict or underpredict the performance of a group? A test may be equally valid for two groups, but unfair because one group's performance is systematically underpredicted in relation to the other's. The traditional concern in fairness inquiries was that tests might unfairly underpredict black achievement, but recent studies suggest that this is not the case. See Cole, *Bias in Testing*, 36 AM. PSYCHOLOGIST 1067, 1069-70 (1981).

Less-discriminatory-alternative analysis assumes that a given test is valid and fair, and asks whether there is any other valid and fair test that will have a less adverse impact.

on the upper level is vital. Although present selection methods are often hard to prove valid, employers and courts may remain convinced that traditional selection devices serve some useful purposes. If inquiry is focused solely on whether an employer's current methods are valid, the conclusion in many cases is likely to be that the commonsense case for validity is strong enough to justify their continued use. It is therefore important to realize that even a valid employment scheme is not justified under title VII if less discriminatory alternatives exist that are likely to be equally valid.

The potential of this approach can be illustrated by considering a law firm's system for selecting associates. Assume that the law firm makes choices among applicants primarily on the basis of law school attended and law school grades. The firm would presumably defend the job-relatedness of its selection system on the ground that there is significant evidence that law school grades measure abilities relevant to future performance as a lawyer. It would advance similar commonsense arguments to justify quality of law school attended as a selection criterion. While these arguments do not establish validity, a court might be tempted to conclude that the firm's selection process is valid, because it is a commonplace that grades and the quality of prior schooling bear some relation to abilities important to future performance as a lawyer and because law firms routinely rely on these criteria.

A less-discriminatory-alternative approach requires the court to consider the likelihood that there are equally valid alternative systems that have a less discriminatory impact. High grades and attendance at prestigious law schools are, at best, imperfect predictors of performance as a lawyer. Grades, for example, are designed to measure abilities that constitute a very small segment of the range of abilities involved in effective lawyering. Alternative selection schemes might measure additional attributes of at least equal importance. The firm might, for example, develop an assessment system that gave far less weight to law school examinations and more to performance on long-term written projects, trial and appellate advocacy skills, ability to work well with colleagues on cooperative projects, and qualities of aggressiveness, energy, and dedication to work.

There are probably many less discriminatory alternatives to many of our traditional selection systems. Over the years, experts have developed numerous alternative selection methods that have a less adverse racial impact and that seem likely to be at least as valid as traditional methods. For example, they have suggested substituting job-sample performance tests for

traditional paper-and-pencil tests as predictors for lower level jobs.²³³ They have substituted tests based on preparation manuals for traditional aptitude tests, and found a reduction in adverse impact. The new tests, supported by content validity studies, were designed to measure only knowledge obtained from materials that candidates would read during their training and probationary periods. The theory is that aptitude tests measure learning skills and have an adverse impact on minorities who have received inadequate training in such skills. Tests based on preparation manuals, on the other hand, allow motivation to influence test performance. Testing experts have also developed assessment center techniques that have been praised as improvements over traditional management-level selection and that have been reported to have, in many instances, a less adverse racial impact.²³⁴

Courts looking at lower level employment systems have felt a distance that enabled them to imagine the possibility of less discriminatory alternatives. They have struck down prior experience requirements, for example, finding that performance tests or on-the-job training programs are less discriminatory alternatives that serve the employer's purposes equally well.²³⁵

In upper level cases, courts must try to distance themselves from the selection systems with which they are most familiar, so that they can be similarly receptive to alternative ways of doing things. Plaintiffs should explore alternative selection devices, rather than focusing simply on the validity of existing devices. And employers should realize that such alternatives present opportunities for potentially useful innovation.²³⁶

There is no question that many upper level selection systems appear to make some sort of sense. There is serious

²³³ See, e.g., Schmidt, Greenthal, Hunter, Berner & Seaton, *Job Sample v. Paper-and-Pencil Trades and Technical Tests: Adverse Impact and Examinee Attitudes*, 30 PERSONNEL PSYCHOLOGY 187 (1977).

²³⁴ See, e.g., Byham, *Assessment Centers for Spotting Future Managers*, HARV. BUS. REV., July-Aug. 1970, at 150.

²³⁵ See, e.g., *Crockett v. Green*, 388 F. Supp. 912 (E.D. Wis. 1975), *aff'd*, 534 F.2d 715 (7th Cir. 1976); see also *Bush v. Lone Star Steel Co.*, 373 F. Supp. 526, 539-40 (E.D. Tex. 1974) (merger of black and white seniority lines ordered, with a provision that the company could test skill and ability by a trial or break-in period of no less than five days).

²³⁶ Educational experts had for decades condemned the civil service testing system challenged in *Chance v. Board of Examiners*, 458 F.2d 1167 (2d Cir. 1972), on the ground that it was largely responsible for perpetuating the mediocre quality of the school system's supervisory staff. Educational experts had also long argued that greater reliance should be placed on assessment of on-the-job performance than on written examinations. As a result of various consent decrees negotiated in *Chance*, on-the-job performance became a major part of the licensing process for the school system's supervisors.

question whether upper level employers have begun to think imaginatively about alternative selection devices that might have a lesser racial impact but serve their purposes as well or better.

V. CONCLUSION

The courts are divided in their approach to the upper level title VII cases that have arisen with increasing frequency during recent years. Some have struggled to apply the doctrines developed in lower level cases — doctrines that give real meaning to title VII. Others have granted upper level employers effective immunity from title VII challenge.

Courts reluctant to subject upper level employers to the heavy burdens of justification that title VII has traditionally imposed should at least face up to the issue whether a differential standard is justified. Rational inquiry is not advanced when courts manipulate procedural and substantive doctrines and pretend not to see that racial exclusion is taking place. Moreover, putting the stamp of judicial approval on racially exclusionary systems is seriously unjust if in fact the “finding” that the systems are job related is based simply on judges’ views that systems with which they are personally familiar make sense.

The argument of this Article is that traditional title VII standards should in fact be applied: employers should be forced to demonstrate the necessity for racially exclusionary policies. This requirement might well lead to the discovery that many traditional selection systems cannot be justified by concerns with high quality performance — that these systems have often been designed to serve a variety of factors having nothing to do with merit.

Such a revelation might force us to consider imaginative alternatives to traditional selection methods — alternatives that we should consider if our only concern is quality, but that we have failed to consider because of the tendency of those who are “in” to perpetuate the systems that got them there.

Another result might be increased pressure for quota or racially proportionate hiring among those satisfying certain minimum qualifications. This pressure would be in large part the consequence of careful scrutiny of upper level employment systems. For it will probably be difficult to demonstrate the job-relatedness and business necessity of many of these selection systems, in large part because it is difficult to determine what good performance is and how to measure it. Employers in cases involving business managers, lawyers, doctors, teach-

ers, and the like will find it difficult to say what they are looking for, beyond a level of minimum competence, and how they know when they have found it. Moreover, value judgments inevitably infect such evaluations. Is the best business manager one who simply maximizes production, or should we be looking for business managers who in addition create satisfying working environments? Is the best school principal one who possess advanced academic degrees, or one who will seek to learn about and to fulfill the needs of the community the school serves? Is the best doctor one who will excel at state-of-the-art surgical procedures, or one who will provide badly needed health care to poor families in unglamorous communities? To the extent that such value judgments are inevitable, to the extent that employers cannot say what makes for the ideal job-incumbent, to the extent that those in power almost inevitably create selection devices and define jobs in ways that are self-serving and self-perpetuating, it is hard to justify selection devices with a racially exclusionary impact. Selection on a racially proportionate basis seems an appropriate solution. Ethnic minorities would then be represented in positions in which they could make the essentially political judgments about how our major social institutions should operate and how jobs carrying significant power to influence society should look.