Considerations in Addressing Adverse Impact

IPMA-HR
INTERNATIONAL PUBLIC MANAGEMENT ASSOCIATION for HUMAN RESOURCES

Considerations in Addressing Adverse Impact

Dianna E. Belman, Author
Toni Kovalski and Richard Geiger, Editors

The International Public Management Association for Human Resources

The International Public Management Association for Human Resources (IPMA-HR) represents human resource professionals in federal, state, and local government.

Since 1953, IPMA-HR's Assessment Services Department has been providing high-quality, reliable test products and services to the public safety community for the purposes of hiring and promotions. With more than 200,000 tests issued annually to federal, state, county, and local municipalities in the United States, Canada and the European Union, IPMA-HR is your trusted source for test products and services.

At IPMA-HR, our customers are our most important partners. We strive to provide quality products and services and efficient customer service; that is our commitment to the profession and to our customers.

If you need additional information, please contact:

Assessment Services Department
1617 Duke Street
Alexandria, Virginia 22314
Phone: (800) 381-TEST (8378) Fax: (703) 684-0948
assessment@ipma-hr.org

First Edition, First Printing, August 2009 Printed in the United States of America

ACKNOWLEDGEMENTS

This is the first edition of Considerations in Addressing Adverse Impact. Acknowledgement and special thanks is extended to **Ms. Toni Kovalski** of TJK Consulting who provided thorough reviews of the rough drafts of this handout.

Special appreciation is also extended to **Mr. Richard Geiger** for his editorial skills.

INTRODUCTION

The International Public Management Association for Human Resources' (IPMA-HR) Assessment Services Department frequently receives inquiries regarding adverse impact—how to calculate it and how to reduce it. This handout was designed to provide IPMA-HR test users with information to consider when addressing adverse impact.

With a few exceptions, the information contained in this handout is largely based on two books, *Adverse Impact and Test Validation* by Dan Biddle (2006) and *Employment Discrimination Litigation* edited by Frank J. Landy (2005).

The information contained in this handout should not be construed as legal advice or opinion. It is the responsibility of the agency to make decisions regarding selection procedures. IPMA-HR strongly recommends that the agency consult with its legal staff regarding selection procedures. Only the jurisdiction is aware of local regulations and legal precedents that should dictate the agency's actions.

HISTORY AND DEVELOPMENT OF ADVERSE IMPACT

Civil Rights Act of 1964

The origins of adverse impact are rooted in the history of discriminatory practices in the United States of America. On July 2, 1964, the 88th Congress of the United States of America passed the Civil Rights Act, a landmark piece of legislation that outlawed discriminatory practices:

> In a nationally televised address on June 6, 1963, President John F. Kennedy urged the nation to take action toward guaranteeing equal treatment of every American regardless of race. Soon after, Kennedy proposed that Congress consider civil rights legislation that would address voting rights, public accommodations, school desegregation, nondiscrimination in federally assisted programs, and more.
>
> Despite Kennedy's assassination in November of 1963, his proposal culminated in the Civil Rights Act of 1964, signed into law by President Lyndon Johnson just a few hours after House approval on July 2, 1964. The act outlawed segregation in businesses such as theaters, restaurants, and hotels. It banned discriminatory practices in employment and ended segregation in public places such as swimming pools, libraries, and public schools. (Retrieved from http://www.ourdocuments.gov/doc.php?doc=97 on 12/05/2008)

Title VII of the Civil Rights Act is the most frequently cited statutory basis for employment discrimination challenges involving cut scores (Kehoe and Olson, 2005) and has four general stipulations regarding the implementation of selection procedures and adverse impact (Mueller, Norris, and Oppler, 2005). The first and second stipulations state that any procedure resulting in adverse impact with respect to a protected subgroup must be valid and job relevant, and it must provide some economic or strategic utility to the organization. The third stipulation states that even if the first and second stipulations are met, a selection procedure may still be illegal if there are alternative procedures that could be used, which do not result in what was later termed

"adverse impact." The fourth stipulation was added in 1991, and it prohibits the use of procedures where scores are adjusted on the basis of race, color, religion, sex, or national origin.

Equal Employment Opportunity Commission and Guidelines

The Civil Rights Act of 1964 also resulted in the creation of the Equal Employment Opportunity Commission (EEOC). On August 24, 1966, two years after the passage of the Civil Rights Act, the term "adverse impact" was introduced in a treatise called the *EEOC Guidelines on Employment Testing Procedures* (hereafter, the *1966 EEOC Guidelines*). The *1966 EEOC Guidelines* may be thought of as the earliest predecessor to the *1978 Federal Uniform Guidelines on Employee Selection Procedures* (hereafter, the *1978 Uniform Guidelines*).

Since the issuance of the *1966 EEOC Guidelines*, the U.S. Federal Government has endeavored to define what adverse impact really means and how it should be calculated, which is why the 1978 Uniform Guidelines has so many predecessors. The *1966 EEOC Guidelines*, for example, were ambiguous with regard to adverse impact and did not indicate how it should be determined if an employer's selection practices have adverse impact (Biddle, 2006).

Beginning in 1968, a series of revisions to the *1966 EEOC Guidelines* were issued. In 1968, the U.S. Department of Labor issued the *Employment Tests by Contractors & Subcontractors*. This piece of legislation required that data must be reported separately for minority and nonminority groups whenever feasible (Lawshe, 1987). In 1970, the *1966 EEOC Guidelines* were revised to include the term "differential validity." Additionally, the *1970 EEOC Guidelines* indicated that instances of higher rejection rates for minority and nonminority candidates may constitute evidence of discrimination (Lawshe, 1987).

Employee Testing and Other Selection Procedures

1971 was a very eventful year for adverse impact. The U.S. Department of Labor issued the *Employee Testing and Other Selection Procedures*, which basically reiterated the *1970 EEOC Guidelines*. That same year, the *Office of Federal Contract Compliance Guidelines* attempted to define discrimination but did not provide specific guidance on how it should be calculated (Biddle, 2006).

***Griggs v. Duke Power Company* (1971)**

On March 8, 1971, the Judicial Branch of the U.S. Government ruled on a landmark case, *Griggs v. Duke Power Company*, which helped to further define adverse impact. "Duke Power was using a high-school diploma and an off-the-shelf intelligence test as screening devices, both of which had adverse impact against blacks. Since the jobs being tested did not appear to really require a high school diploma to be performed successfully, the court held that the employer had to show a 'business necessity' for these two requirements, or Duke Power would be in violation of Title VII. Responding to the adverse impact that these two requirements had on blacks, the court only stated a few words: '…they operated to disqualify blacks at a *substantially higher rate* than white applicants.' Once again, the Equal Employment Opportunity (EEO) had only a 'what' but not a 'how' regarding adverse impact. (Biddle, 2006, p. 2).

Technical Advisory Committee on Testing

In 1971, the State of California Fair Employment Practice Commission (FEPC) created an advisory committee called the Technical Advisory Committee on Testing (TACT). Composed of 32 specialists from various labor, employment, and technical fields, TACT was charged with deliberating on the techniques that should be used to evaluate adverse impact (Biddle, 2006).

TACT's deliberations led to the development of the "Four-Fifths Rule." The Four-Fifths Rule is calculated by dividing the passing rate of the group with the lowest selection rate (usually minority group members and females) by the passing rate of the group with the highest selection rate (usually white group members or males). Any value less than

four-fifths (i.e., 80%) is said to violate the threshold test for evaluating adverse impact. For more on calculating adverse impact, see *Adverse Impact Analyses*.

The Four-Fifths Rule (i.e., the 80% Rule)

Given the prevalence of the Four-Fifths Rule, it is important to understand the logic behind its development. According to Biddle (2006), the TACT committee members wanted to propose that a statistical test be used to calculate adverse impact. However, they also agreed that the individuals responsible for calculating adverse impact would most likely not have the appropriate knowledge, skills, and abilities necessary to manually perform statistical tests[1]. After much deliberation, the TACT committee members came up with a compromise. They agreed that they needed an administrative guideline as well as a technical guideline in the event of legal challenges. A member of the TACT committee recalls how the Four-Fifths Rule was created:

> I recall a heated debate that went on for way too long (as usual) with two camps: a 70% camp and a 90% camp. The 80% Test was born out of two compromises: (1) a desire expressed by those writing and having input into the Guidelines to include a statistical test as the primary step but knowing from an administrative point of view a statistical test was not possible for the FEPC consultants who had to work the enforcement of the Guidelines, and (2) a way to split the middle between the two camps, the 70% camp and the 90% camp. A way was found to use both. In the way the 80% Test was defined by TACT, if there was no violation of the 80% Test, then there would be no reason to apply statistical significance tests. This hopefully would eliminate many calculations and many situations where TACT would not be necessary and the decision could be made in the field. So from the practical point of view, the 80% Test became a first step. If there was no 80% Test violation, there was no need to go further and use a statistical test. If there was a violation of the 80% Test, statistical significance

[1] Prior to the frequent use of computers, probability statistics were calculated by hand and, therefore, only calculated by the technically savvy (Biddle, 2005).

> was needed and the 80% Test then became a practical significance test for adverse impact. (Biddle, 2006, p. 3)

In addition to the creation of the Four-Fifths Rule, TACT's deliberations resulted in the *1972 State of California Guidelines on Employee Selection Procedures*. "Some content from these California Guidelines were designed to 'supersede and enlarge upon' (Section 3 preamble) the earlier set of Guidelines on Employment Testing Procedures, issued by the US EEOC on August 24, 1966. Some content from these California Guidelines was later incorporated into the Federal Uniform Guidelines on Employee Selection Procedures (1978), a document still in force at the time of writing" (Biddle, 2006, p. 3).

State of California Guidelines on Employee Selection Procedures (1972)

The *1972 State of California Guidelines on Employee Selection Procedures* contained the first concrete definition of adverse effect. As a forerunner of adverse impact, the definition of adverse effect foreshadows the idea that adverse impact should be based on comparisons of selection rates between applicant groups:

> Adverse effect refers to a total employment process which results in a significantly higher percentage of a protected group in the candidate population being rejected for employment, placement, or promotion. The difference between the rejection rates for a protected group and the remaining group must be statistically significant at the .05 level. In addition, if the acceptance rate of the protected group is greater than or equal to 80% of the acceptance rate of the remaining group, then adverse effect is said to be not present by definition (Section 7.1). (Biddle, 2006, p. 4).

Federal Executive Agency Guidelines on Employee Selection Procedures (1976)

In 1976, the U.S. Department of Justice issued the *Federal Executive Agency Guidelines on Employee Selection Procedures*. Noteworthy contributions of this issuance included eliminating the notion of differential validity from the guidelines and

modifying the term “adverse impact” to include verbiage about fairness and differential rates of selection.

Uniform Guidelines on Employee Selection Procedures (1978)

On August 25, 1978, the Equal Employment Opportunity Commission (EEOC), the Civil Service Commission (CSC), the Department of Labor (DOL), and the Department of Justice (DOJ) adopted the *Uniform Guidelines on Employee Selection Procedures* (henceforth, the *1978 Uniform Guidelines*). “The guidelines are designed to aid in the achievement of our nation's goal of equal employment opportunity without discrimination on the grounds of race, color, sex, religion or national origin. The Federal agencies have adopted the Guidelines to provide a uniform set of principles governing use of employee selection procedures which is consistent with applicable legal standards and validation standards generally accepted by the psychological profession and which the Government will apply in the discharge of its responsibilities” (EEOC, CSC, DOL, and DOJ, 1978, Q&A, I. Purpose and Scope).

In general terms, the *1978 Uniform Guidelines* defines adverse impact as, “A substantially different rate of selection in hiring, promotion, or other employment decision which works to the disadvantage of members of a race, sex, or ethnic group” (EEOC, CSC, DOL, and DOJ, 1978, Section 16: Definitions B). More specifically, the *1978 Uniform Guidelines* defines adverse impact as:

> A selection rate for any race, sex, or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact. Smaller differences in selection rate may nevertheless constitute adverse impact, where they are significant in both statistical and practical terms or where a user's actions have discouraged applicants disproportionately on grounds of race, sex, or ethnic group. Greater differences in selection rate may not constitute adverse impact where the differences are based on small numbers and are not

statistically significant, or where special recruiting or other programs cause the pool of minority or female candidates to be atypical of the normal pool of applicants from that group. Where the user's evidence concerning the impact of a selection procedure indicates adverse impact but is based upon numbers which are too small to be reliable, evidence concerning the impact of the procedure over a longer period of time and/or evidence concerning the impact which the selection procedure had when used in the same manner in similar circumstances elsewhere may be considered in determining adverse impact. Where the user has not maintained data on adverse impact as required by the documentation section of applicable guidelines, the Federal enforcement agencies may draw an inference of adverse impact of the selection process from the failure of the user to maintain such data, if the user has an underutilization of a group in the job category, as compared to the group's representation in the relevant labor market or, in the case of jobs filled from within, the applicable work force. (EEOC, CSC, DOL, and DOJ, 1978, Section 4D).

Given the complexity of the *1978 Uniform Guidelines*' definition of adverse impact, it is important to separate the components and examine them piece by piece.

1. Adverse impact exists in any employment situation in which the selection or hiring rate for any race, sex, or ethnic group[2] is:
    - Less than 80% of the selection rate for the applicant group with the highest selection rate.
    - Close to but not quite 80% of the selection rate for the applicant group with the highest selection rate if the difference in the selection rates is statistically significant (i.e., unlikely due to chance) and practically significant.
2. Adverse impact may exist where evidence concerning the impact of a particular selection procedure indicates that adverse impact exists regardless of the following:

[2] It is important to note that the definition of a minority group member is changing. In fact, the OFFCP anticipates that some discrimination cases may be filed on behalf of white applicants within the next 10 years.

- The evidence is based upon applicant pools that are too small to be reliable.
- The evidence consists of data on the selection procedure that was gathered over a long period of time.

- The only evidence that exists concerning the impact of the selection procedure comes from another user who used the procedure in the same way and under similar circumstances.

- Even if the user has not collected data on adverse impact, Federal enforcement agencies may infer the impact of a selection procedure based on the group's representation in the relevant labor market or the user's workforce (in the case of jobs filled internally).

3. Adverse impact **may not exist** in employment situations where:
    - Differences in the selection rates are based on small, unreliable numbers of applicants that are not statistically significant.

    - There are special recruiting programs in place that result in an atypical pool of minority or female candidates.

Uniform Employee Selection Guidelines Interpretation and Clarification (Questions and Answers) (1980)

On May 2, 1980, the same consortium that created the *1978 Uniform Guidelines* issued the *Uniform Employee Selection Guidelines Interpretation and Clarification (Questions and Answers)*, which is a series of questions and answers designed to clarify and interpret, but not to modify, the *1978 Uniform Guidelines*.

EMPLOYMENT DISCRIMINATION LITIGATION

This section provides an overview of relevant Supreme Court cases pertaining to adverse impact.

Phases of Employment Litigation

Before discussing the cases, it is important to have a general understanding of the steps involved in a typical employment discrimination case:

1. An individual contacts an administrative agency such as the EEOC to register the protest formally.
    a. The administrative agency then investigates the charge of discrimination by requesting basic information from the individual's employer about the charge.
    b. If the administrative agency determines that the charge has merit, the agency will attempt to reach an amicable resolution, called conciliation.

2. The charging party receives a right to sue notification:
    a. The charging party may request this notification after a set time period has elapsed, and the EEOC is required to comply.
    b. The notification will be automatically issued if the EEOC finds the charge without merit.
    c. The notification will be issued if the EEOC finds the charge with merit but cannot resolve the charge with the employer.

3. If the charging party decides to file a lawsuit, then he or she must file a formal complaint with the court. The complaint indicates the name(s) of the plaintiff(s), the reason for the complaint, the right the plaintiff has to file a charge, the alleged discriminatory practice; and the type of resolution sought by the plaintiff.

4. If the plaintiff decides to sue his or her employer as part of a group of individuals who are similarly situated, then the plaintiff's counsel would propose

the group as a class. This step is known as class certification. A judge must decide whether the lawsuit may proceed on behalf of a class, using a set of established criteria including, but not limited to, whether the plaintiffs are all members of a protected class and if there is a common basis for the complaints.

5. Lawyers for both the plaintiff and defendant begin building their cases by gathering information relevant to the charge of discrimination. This is known as the discovery process. Important aspects of the discovery process include gathering information such as documents describing the development of the selection device or procedure, sending interrogatories, and taking depositions. Interrogatories refer to a formal set of written questions submitted by one lawyer to opposing counsel in order to clarify matters of evidence and help determine in advance what facts will be presented at trial. A deposition is a formal interview taken under oath conducted by opposing counsel.

6. Throughout the case, counsel will file various motions, which ask the judge to take certain actions such as requesting information from opposing counsel. There are certain specific types of motions. *Daubert* motions, for example, refer to attempts to rule out expert testimony on the grounds that the testimony is lacking in technical details and may mislead the judge or the jury.

7. Due to the costs and risks associated with trials, most cases do not go to trial and end up being settled out of court. Settlement discussions can begin at any point, but the discussions may become more serious as the trial approaches or once the discovery stage has concluded.

8. If the case goes to trial, it may take place in front of a judge (called a bench trial) or a jury (called a jury trial), although most employment discrimination cases involve a jury. Both parties present their cases and follow with rebuttal. Sometimes the judge asks the plaintiffs to present evidence of adverse impact, which the defendants then rebut. “There is often a third phase of the trial as

well—the consideration of *alternatives*. If the plaintiffs demonstrate adverse impact for a device or procedure and the defendants demonstrate the job-relatedness of that device or procedure, the plaintiffs still have the opportunity to demonstrate that there was an alternative device or procedure available for the defendants for making the personnel decision, an alternative that had equal validity and less adverse impact" (Gutman, 2005, p. 15).

9. Two types of witnesses are called upon during an employment discrimination trial—fact witnesses and expert witnesses. Fact witnesses provide information about factual issues concerning the case, such as describing their personal experiences with the selection process. A specific type of fact witness, whose testimony carries substantial weight, is a 30(b)6 witness who may testify as an official spokesperson for the company about the history and implementation of the company's policies and procedures. Expert witnesses offer opinions about technical issues concerning the case, such as the validity of a selection device or procedure. "In employment discrimination cases, the typical expert witnesses include statisticians who consider issues related to adverse impact determination, economists who consider issues of monetary damages, social psychologists who discuss issues of stereotyping, and I-O [industrial-organizational] psychologists who consider HR and psychometric issues" (Gutman, 2005, p. 17).

10. Once all the evidence has been presented, the judge or jury renders a decision. "When the decision is in favor of the plaintiffs, there is usually a monetary award of some amount. In addition to a monetary award a judge may also order changes in procedures or practices. When the defendants win, the judge may order the plaintiffs to pay some portion of the costs incurred by the defendants in defending against the charge, but these costs are usually minimal and only cover administrative expenses such as copying, fees for court reports, and limited travel expenses.

11. Appeals.

For more information on the steps involved in employment litigation, please see Landy (2005).

According to Gutman (2005), the Supreme Court has made eight major rulings with regard to adverse impact—four of which involved standardized cognitive tests. These four rulings include *Griggs* v. *Duke Power* (1971), *Albemarle Paper Co.* v. *Moody* (1975), *Washington* v. *Davis* (1976), and *Connecticut* v. *Teal* (1982).

***Griggs v. Duke Power* (1971)**

In *Griggs* v. *Duke Power* (1971), Black employees at Duke Power Company filed a lawsuit pursuant to Title VII of the Civil Rights Act of 1964, challenging the requirement that applicants must have a high school diploma or pass an intelligence test as a prerequisite to hire into any of the four higher levels or promotion to the upper three levels (Gutman, 2005). "Incidentally, the diploma requirement was not new. It was instituted in 1955 for white workers seeking promotion to upper-level jobs. Interestingly, there was no evidence that upper-level whites with diplomas performed any better than upper-level whites without diplomas" (Gutman, 2005, p. 27). While the Civil Rights Act of 1964 makes it an unlawful employment practice for an employer to limit, segregate, or classify employees to deprive them of employment opportunities or adversely to affect their status because of race, color, religion, sex, or national origin, it also authorizes the use of any professionally developed ability test, provided that it is not designed, intended, or used to discriminate. The Supreme Court held the following:

> 1. The [Civil Rights] Act requires the elimination of artificial, arbitrary, and unnecessary barriers to employment that operate invidiously to discriminate on the basis of race, and, if, as here, an employment practice that operates to exclude Negroes cannot be shown to be related to job performance, it is prohibited, notwithstanding the employer's lack of discriminatory intent. Pp. 429-433.

2. The [Civil Rights] Act does not preclude the use of testing or measuring procedures, but it does proscribe giving them controlling force unless [401 U.S. 424, 425] they are demonstrably a reasonable measure of job performance. Pp. 433-436.

(Retrieved from http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=401&invol=424 02/09/2009)

***Albemarle Paper Co. v. Moody* (1975)**

In *Albemarle Paper Co.* v. *Moody* (1975), a certified class of present and former black employees filed a lawsuit against both their employer, Albemarle Paper Company, and union, seeking injunctive relief against "any policy, practice, custom or usage" at the Albemarle Paper Company that violates Title VII of the Civil Rights Act of 1964, as amended by the Equal Employment Opportunity Act of 1972. The major issues discussed during the trial were the plant's seniority system, its program of employment testing, and backpay:

> The District Court found that, following a reorganization under a new collective-bargaining agreement, the Negro employees had been "'locked' in the lower paying job classifications," and ordered petitioners to implement a system of plantwide seniority. The court refused, however, to order backpay for losses sustained by the plaintiff class under the discriminatory system, on the grounds that (1) Albemarle's breach of Title VII was found not to have been in "bad faith," and (2) respondents, who had initially disclaimed interest in backpay, had delayed making their backpay claim until five years after the complaint was filed, thereby prejudicing petitioners. The court also refused to enjoin or limit Albemarle's testing program, which respondents had contended had a disproportionate adverse impact on blacks and was not shown to be related to job performance, the court concluding that "personnel tests administered at the plant have undergone validation studies and have been proven to be job related." Respondents appealed on the backpay and pre-employment tests issues. The Court of Appeals reversed the District Court's judgment.

Held:

1. Given a finding of unlawful discrimination, backpay should be denied only for reasons that, if applied generally, would not frustrate the central statutory purposes manifested by Congress in enacting Title VII of eradicating discrimination throughout the [422 U.S. 405, 406] economy and making persons whole for injuries suffered through past discrimination. Pp. 413-422.

2. The absence of bad faith is not a sufficient reason for denying backpay, Title VII not being concerned with the employer's "good intent or absence of discriminatory intent," for "Congress directed the thrust of the Act to the consequences of employment practices, not simply the motivation," Griggs v. Duke Power Co., 401 U.S. 424, 432 . Pp. 422-423.

3. Whether respondents' tardiness and inconsistency in making their backpay demand were excusable and whether they actually prejudiced petitioners are matters that will be open to review by the Court of Appeals if the District Court, on remand, decides again to decline a backpay award. Pp. 423-425.

4. As is clear from Griggs, supra, and the Equal Employment Opportunity Commission's Guidelines for employers seeking to determine through professional validation studies whether employment tests are job related, such tests are impermissible unless shown, by professionally acceptable methods, to be "predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated." Measured against that standard, Albemarle's validation study is materially defective in that (1) it would not, because of the odd patchwork of results from its application, have "validated" the two general ability tests used by Albemarle for all the skilled lines of progression for which the two tests are, apparently, now required; (2) it compared test scores with subjective supervisorial rankings, affording no means of knowing what job-performance criteria the supervisors were considering; (3) it focused mostly on job groups

near the top of various lines of progression, but the fact that the best of those employees working near the top of a line of progression score well on a test does not necessarily mean that the test permissibly measures the qualifications of new workers entering lower level jobs; and (4) it dealt only with job-experienced, white workers, but the tests themselves are given to new job applicants, who are younger, largely inexperienced, and in many instances nonwhite. Pp. 425-435.

5. In view of the facts that during the appellate stages of this litigation Albemarle has apparently been amending its departmental organization and the use made of its tests; that issues of standards of proof for job relatedness and of evidentiary procedures involving validation tests have not until now been clarified; [422 U.S. 405, 407] and that provisional use of tests pending new validation efforts may be authorized, the District Court on remand should initially fashion the necessary relief. P. 436.
(Retrieved from http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=422&invol=405 on 02/09/2009)

***Washington v. Davis* (1976)**

In *Washington* v. *Davis* (1976), two Black candidates, who had previously applied for the position of police officer, filed a claim that Test 21[3] was not job-related and had adverse impact for black applicants. Both parties filed cross-motions for summary judgment. In siding with the plaintiff, the District Court acknowledged that the number of black police officers is not proportionate to the City's demographics, a higher percentage of black candidates fail Test 21 than white candidates, and Test 21 had not been validated. Despite this acknowledgment, the District Court concluded that the plaintiffs were not entitled to relief and granted the defendant's motion for summary judgment given the following: (a) 44% of new police recruits were black—a percentage that was in proportion to the number of black police officers who served on the police

[3] Test 21 refers to the name of a particular examination administered to prospective government employees in order to assess their verbal skills.

force and to the number of blacks aged 20 to 29 years old in the recruiting area, (b) the Washington, D.C., Police Department sought to recruit black candidates, many of whom passed the test but failed to report for duty, and (c) Test 21 was a useful indicator of training school performance, which precluded the need to demonstrate validity. The test was not designed to, and did not, discriminate against otherwise qualified black candidates. The plaintiffs filed an appeal on the grounds that Test 21 has adverse impact for black candidates. The Court of Appeals reversed the ruling:

> The Court of Appeals reversed, and directed summary judgment in favor of respondents, having applied to the constitutional issue the statutory standards enunciated in *Griggs v. Duke Power Co.*, 401 U.S. 424 , which held that Title VII of the Civil Rights Act of 1964, as amended, prohibits the use of tests that operate to exclude members of minority groups, unless the employer demonstrates that the procedures are substantially related to job performance. The court held that the lack of discriminatory intent in the enactment and administration of Test 21 was irrelevant; that the critical fact was that four times as many blacks as whites failed the test; and that such disproportionate impact sufficed to establish a constitutional violation, absent any proof by petitioners that the test adequately measured job performance. Held:
>
> 1. The Court of Appeals erred in resolving the Fifth Amendment issue by applying standards applicable to Title VII cases. Pp. 238-248.
>
> (a) Though the Due Process Clause of the Fifth Amendment contains an equal protection component prohibiting the Government from invidious discrimination, it does not follow that a law or other official act is unconstitutional solely because it has a racially disproportionate impact regardless of whether it reflects a racially discriminatory purpose. Pp. 239-245.
>
> (b) The Constitution does not prevent the Government from seeking through Test 21 modestly to upgrade the communicative abilities of its employees rather than to be satisfied with some lower level of competence, particularly where the job

requires special abilities to communicate orally and in writing; and respondents, as Negroes, could no more ascribe their failure to pass the test to denial of equal protection than could whites who also failed. Pp. 245-246.

(c) The disproportionate impact of Test 21, which is neutral on its face, does not warrant the conclusion that the test was a purposely discriminatory device, and on the facts before it the District Court properly held that any inference of discrimination was unwarranted. P. 246.

(d) The rigorous statutory standard of Title VII involves a more probing judicial review of, and less deference to, the seemingly reasonable acts of administrators and executives than is [426 U.S. 229, 231] appropriate under the Constitution where, as in this case, special racial impact but no discriminatory purpose is claimed. Any extension of that statutory standard should await legislative prescription. Pp. 246-248.

2. Statutory standards similar to those obtained under Title VII were also satisfied here. The District Court's conclusion that Test 21 was directly related to the requirements of the police training program and that a positive relationship between the test and that program was sufficient to validate the test (wholly aside from its possible relationship to actual performance as a police officer) is fully supported on the record in this case, and no remand to establish further validation is appropriate. Pp. 248-252.
(Retrieved from
http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=426&invol=229
on 02/09/2009)

***Connecticut v. Teal* (1982)**

In *Connecticut* v. *Teal* (1982), black employees of a Connecticut state agency were temporarily promoted to supervisory positions. In order to achieve permanent supervisory status, employees had to participate in a selection process. First,

employees had to pass a written test, which was administered to 48 black and 259 white applicants. Fifty-four percent of the black applicants passed the examination, which is approximately 68 percent of the passing rate for the white applicants. Black employees who failed the test were excluded from further consideration for the permanent supervisory positions and, in turn, filed a law suit with the Federal District Court against petitioners (i.e., the State of Connecticut and certain state agencies and officials), alleging that Title VII of the Civil Rights Act of 1964 had been violated, because the promotion process required that applicants pass a written examination that disproportionately excluded blacks and was not job-related. Before the trial, petitioners made promotions from the eligibility list. Approximately 23% of the black candidates and 13.5% of the white candidates were promoted. Petitioners argued that the "bottom-line" result was more favorable to black applicants than it was to white applicants. The District Court agreed, holding that the "bottom line" percentages precluded the finding of a Title VII violation and that petitioners were not required to show that the examination was job-related. The Court of Appeals reversed the decision, holding that the District Court erred in its ruling, and that the results of the examination alone were insufficient to support a prima facie case of disparate impact in violation of Title VII. The Supreme Court held that:

> Petitioners' nondiscriminatory "bottom line" does not preclude respondents from establishing a prima facie case nor does it provide petitioners with a defense to such a case. Pp. 445-456.
>
> (a) Despite petitioners' nondiscriminatory "bottom line," respondents' claim of disparate impact from the examination, a pass-fail barrier to employment opportunity, states a prima facie case of employment discrimination under 703(a)(2) of Title VII, which makes it an unlawful employment practice for an employer to "limit, segregate, or classify his employees" in any way which would deprive "any individual of employment [457 U.S. 440, 441] opportunities" because of race, color, religion, sex, or national origin. To measure disparate impact only at the "bottom line" ignores the fact that Title VII guarantees these individual black respondents the opportunity to compete equally with white

workers on the basis of job-related criteria. Respondents' rights under 703(a)(2) have been violated unless petitioners can demonstrate that the examination in question was not an artificial, arbitrary, or unnecessary barrier but measured skills related to effective performance as a supervisor. Pp. 445-451.

(b) No special haven for discriminatory tests is offered by 703(h) of Title VII, which provides that it shall not be an unlawful employment practice for an employer to act upon results of an ability test if such test is "not designed, intended, or used to discriminate" because of race, color, religion, sex, or national origin. A non-job-related test that has a disparate impact and is used to "limit" or "classify" employees is "used to discriminate" within the meaning of Title VII, whether or not it was "designed or intended" to have this effect and despite an employer's efforts to compensate for its discriminatory effect. Pp. 451-452.

(c) The principal focus of 703(a)(2) is the protection of the individual employee, rather than the protection of the minority group as a whole. To suggest that the "bottom line" may be a defense to a claim of discrimination against an individual employee confuses unlawful discrimination with discriminatory intent. Resolution of the factual question of intent is not what is at issue in this case, but rather petitioners seek to justify discrimination against the black respondents on the basis of petitioners' favorable treatment of other members of these respondents' racial group. Congress never intended to give an employer license to discriminate against some employees on the basis of race or sex merely because he favorably treats other members of the employees' group. Pp. 452-456. (Retrieved from http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=457&invol=440 on 04/08/2009)

Civil Rights Act of 1991

On November 22, 1991, the 102nd Congress of the United States of America passed the Civil Rights Act of 1991. This piece of legislation was passed, "[t]o amend the Civil Rights Act of 1964 to strengthen and improve Federal civil rights laws, to provide for damages in cases of intentional employment discrimination, to clarify provisions regarding disparate impact actions, and for other purposes." The Civil Rights Act of 1991 was passed in response to a series of United States Supreme Court decisions that restricted the rights of employees who had sued their employers on grounds of discrimination. These cases include *Patterson v. McLean Credit Union* (1989), *Wards Cove Packing Co. v. Atonio* (1989), *Price Waterhouse v. Hopkins* (1989), and *Martin v. Wilks* (1989).

***Patterson v. McLean Credit Union* (1989)**

In *Patterson v. McLean Credit Union* (1989), a black female was laid off by McLean Credit Union where she worked as a teller and file coordinator for 10 years. The petitioner, Patterson, brought this action in District Court under 42 U.S.C. § 1981, alleging that the respondent, McLean Credit Union, harassed her, did not promote her, and discharged her, all due to her race. However, the District Court determined that a claim for racial harassment is not actionable under § 1981 and declined to submit that part of the case to the jury. Instead, the District Court informed the jury that, in order for the petitioner to successfully argue that she was not promoted because of her race, she must demonstrate that she was better qualified than the white employee who had purportedly been promoted. The jury ruled in favor of the respondent regarding both the promotion discrimination claim and discriminatory discharge claim. The Court of Appeals concurred with the District Court ruling in favor of the respondent.

The Supreme Court held that the right to make contracts does not extend to conduct by an employer after the contractual relation has been established, including breach of the terms of the contract or imposition of discriminatory working conditions. Essentially, the Supreme Court's decision in *Patterson v. McLean Credit Union (1989)* restricted a decision made in an earlier case, *Runyon v. McCrary* (1976).

Patterson formally involved the question of whether an African-American woman's claim of racial harassment in employment stated a cause of action under Title 42, section 1981, of the U.S. Code, a surviving portion of the Civil Rights Act of 1866. In Runyon v. McCrary (1976) and Jones v. Alfred H. Mayer Co. (1968), section 1981 and a companion provision had been interpreted to reach private racial discrimination in contractual and property relations. After the initial argument, the Court, on its own motion, ordered a reargument and requested that the parties address the question of whether *Runyon*'s interpretation of section 1981 should be overruled. Patterson thus seemed on the verge of becoming a landmark case reversing the prior twenty years' practice of applying the 1866 act's modern counterparts to cases involving private discrimination.

Few procedural orders in the Supreme Court's history have caused such a volatile reaction. Within the Court, the reargument order itself prompted sharp dissents from Justices Harry Blackmun and John Paul Stevens, both joined by Justices William Brennan and Thurgood Marshall. These dissents moved the majority to take the unusual steps of defending a reargument order in writing. The civil rights community, the press, and scholarly journals focused intense attention on the pending case.

After the second argument, relying on the doctrine of *stare decisis* (see Precedent), the Court unanimously declined to overrule *Runyon*. But, in an unprecedented interpretation of section 1981 that prompted four dissents, the Court held that the right to make contracts does not extend to conduct by an employer after establishment of the contractual relation, including Patterson's claim of posthiring racial harassment. Congress reacted to Patterson and other decisions by passing the Civil Rights Act of 1991, which overruled Patterson's narrow reading of section 1981. (http://www.encyclopedia.com/doc/1O184-PattersonvMcleanCreditUnn.html retrieved on 05/12/2009)

***Wards Cove Packing Co. v. Atonio* (1989)**

In *Wards Cove Packing Co. v. Atonio* (1989), a class of nonwhite workers filed suit in the District Court under Title VII of the Civil Rights Act of 1964 that the selection procedures used by Wards Cove Packing Company, the owner of Alaskan salmon canneries, resulted in racial discrimination. Specifically, the plaintiffs claimed that the majority of skilled or noncannery jobs are filled by white workers, while the majority of unskilled or cannery jobs are filled by minority workers. The District Court found that minority workers were overrepresented in cannery jobs because many of those jobs were filled under a hiring-hall agreement with a union in which the majority of members were not white. For this and other reasons, the District Court rejected the respondent's claim. "The Court of Appeals ultimately reversed in pertinent part, holding, inter alia, that respondents had made out a prima facie case of disparate impact in hiring for both skilled and unskilled noncannery jobs, relying solely on respondents' statistics showing a high percentage of nonwhite workers in cannery jobs and a low percentage of such workers in noncannery positions. The [Court of Appeals] also concluded that once a plaintiff class has shown disparate impact caused by specific, identifiable employment practices or criteria, the burden shifts to the employer to prove the challenged practice's business necessity"
(http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=490&invol=642 retrieved on 05/12/2009).

The Supreme Court ruled that a prima facie case of disparate impact cannot be based solely on the finding that a racial imbalance exists in one part of an employer's workforce, unless there are barriers or practices that deter qualified minority group members from applying for noncannery positions. Additionally, the Court ruled that comparisons by race should be made between individuals in the target job classification and qualified individuals in the relevant labor market.

***Price Waterhouse v. Hopkins* (1989)**

The respondent, Ann Hopkins, was a senior manager at Price Waterhouse when she was proposed for partnership in 1982. Hopkins was neither offered nor denied partnership, but instead her candidacy was held for reconsideration the following year. When the partners in Hopkins' office later refused to reconsider her for partnership, she sued in Federal District Court under Title VII of the Civil Rights Act of 1964. Hopkins charged that her employer had discriminated against her on the basis of sex. "The District Court ruled in respondent's favor on the question of liability, holding that petitioner had unlawfully discriminated against her on the basis of sex by consciously giving credence and effect to partners' comments about her that resulted from sex stereotyping. The Court of Appeals affirmed. Both courts held that an employer who has allowed a discriminatory motive to play a part in an employment decision must prove by clear and convincing evidence that it would have made the same decision in the absence of discrimination, and that petitioner had not carried this burden" (retrieved from http://supreme.justia.com/us/490/228/case.html on 05/26/2009). In *Price Waterhouse v. Hopkins* (1989), the Court shifted the burden of proof, holding that a disparate treatment plaintiff must demonstrate by direct evidence that employment practices substantially depended on illegitimate criteria. The Civil Rights Act of 1991 altered the burden once again.

***Martin v. Wilks* (1989)**

In *Martin v. Wilks (1989)*, a class of African American firefighters filed suit under Title VII of the Civil Rights Act of 1964 alleging that the selection and promotion procedures used by Birmingham, Alabama and a county personnel board resulted in racial discrimination. Prior to approving consent decrees related to goals for the hiring and promotion of African-American firefighters, a federal district court ordered public notice of a hearing on the fairness of the decrees. Representatives from the Birmingham Firefighters Association attended the hearing and objected to the decrees. The Court denied the motion. At the same time, a group of white firefighters brought a reverse discrimination action against the City of Birmingham and the personnel board, arguing

that the consent decree made them victims of racial discrimination because they were being denied promotions in favor of less-qualified candidates.

> The Supreme Court posed the issue as being a choice between whether the African-American plaintiffs should have joined all possibly affected parties before entry of the consent decree or whether possibly affected parties should have to seek to intervene in the lawsuit resulting in the consent decree. The Court held that the white firefighters, not having been parties to the original litigation, were not bound by the consent decree. The Court indicated that plaintiffs who seek to alter employment practices are best able to bear the burden of identifying those who might be adversely affected if plaintiffs prevail. Plaintiffs should join such parties to their lawsuit. Such parties need not seek to intervene. The Civil Rights Act of 1991 limited the scope of Martin v. Wilks and made it marginally more difficult for third parties to attack hiring and promotion decisions based on employment discrimination consent decrees.
> (retrieved from http://www.answers.com/topic/martin-v-wilks on 05/27/2009)

CAUSES OF ADVERSE IMPACT

IPMA-HR's Assessment Services Department frequently receives inquiries from customers asking if our tests have adverse impact. Unfortunately, this is not a question with a clear answer. Guion (Biddle, 2006) indicates that there are many reasons why adverse impact occurs in particular testing situations. Adverse impact can result from chance, the nature of test use, measurement problems inherent to the test, differences in distribution sizes, reliable subgroup differences in general approaches to test taking, and/or true population differences in distributions of the traits being measured.

Interestingly enough, adverse impact is *rarely* caused by inherent bias in tests. "Adverse impact has become a loaded term, fraught with suggestions of ill intent on the part of the employer. It should, however, be noted that adverse impact simply describes differences between groups on a testing process. It is not a legal term that implies guilt or a psychometric term that implies unfairness or test bias. Many employers that test for relevant job skills will generate adverse impact in a testing process in one way or another, and most studies show that adverse impact is not normally due to forms of bias inherent to tests (Sackett, 2001; Neisser, 1996)" (Biddle, 2006, p. 2).

Nevertheless, a common misconception is that adverse impact is caused by the particular selection procedure used in the employment process, and that adverse impact is a property of the selection procedure. "It is important to recognize that adverse impact is not a property of a test; rather, it is an analysis of the results of employment decisions which may be influenced, in part, by a test or any number of other factors (e.g., number of vacancies, seniority, performance ratings, budgetary constraints, etc.). Adverse impact could result from the use of unfair, biased, discriminatory, or unlawful procedures. However, adverse impact could also result from true differences between two protected groups on a relevant, job-related characteristic. In absence of further information, adverse impact should be regarded as a neutral term (retrieved from www.adverseimpact.org on 12/05/2008)."

According to Biddle (2006), “it is not uncommon for public sector employees to have one or more entry-level positions with adverse impact in the testing process. Written tests typically have the highest degree of adverse impact, with the highest level of impact against blacks, then Hispanics, and sometimes Asians (Sackett, 2001; Neisser, 1996). Physical ability tests typically have adverse impact against women, especially when they measure upper body strength” (p. 1).

For entry-level tests, IPMA-HR consultants conduct a fairness analysis which consists of examining incumbent test scores by race/ethnicity and sex. For promotional tests, IPMA-HR’s consultants always have subject matter experts review test content for fairness in terms of whether there is bias in the wording.

EMPLOYER OBLIGATIONS AND BEST PRACTICES

Evaluating the Test

IPMA-HR's Assessment Services Department serves as a test publisher. This means that we provide tests to agencies for the purposes of hiring and promoting. As the test user, it is your obligation to determine if our tests meet your needs. Specifically, when selecting a test from our catalog, you should review the following:

- An inspection copy of the test itself and any additional booklets that accompany the test. Some of our entry-level tests include additional booklets, such as Test Information Packets (TIPs) and scratch paper booklets.
- Any preparatory materials associated with the test, including candidate reading lists and study guides.
- The test response data report associated with the test. A report presenting test data gathered on candidates from jurisdictions that use IPMA-HR tests. This includes information about passing points, score distributions, and adverse impact.
- The technical report associated with the test. A report that provides information on how the test was developed.

With this information in hand, you will be able to make sound decisions concerning test administration. If you want to know if a test has adverse impact, for example, you would want to know whether it has had adverse impact in the situations where it has been administered. To accomplish this, you would want to review the test response data report associated with the test you are thinking of administering. This report presents anonymous test scores earned by candidates by race and gender and uses the 80% rule to calculate adverse impact. Furthermore, you could review the fairness analysis included in the technical report associated with the test in order to determine whether the test predicts performance equally for all protected groups. The report will also help you to determine whether the demographic makeup of the sample used in the validation study is representative of your applicant pool and workforce.

As stated above, a technical report provides information on how the test was developed. Ford and Blair (2006) suggest asking the following questions as you evaluate validation studies:[4]

1. **Does the test have a complete and adequate technical report?**
2. Has the job changed in any important way since the job analysis/validation study was conducted?
3. Were the problems or goals for the validation study relevant to our organization and position?
4. **Did the validation study include a job analysis?**
5. **Is the recommended use of the test consistent with our process?**

All of IPMA-HR's *entry-level* tests are validated using *criterion-related* validation processes. As you read through entry-level technical reports, here are 12 questions to consider when evaluating criterion-related validation evidence (Ford & Blair, 2006): [5]

1. Where has the test been validated?
2. **Has the test been validated in multiple samples?**
3. Was the test development based on the validation sample? If so, you should expect the validity coefficient to be lower in future samples due to capitalization on chance.
4. **Is the validation sample representative of our agency and labor market?**
5. **What criterion was used in validating the test?**
6. Was the criterion appropriately observed, recorded, quantified, and evaluated?
7. Does the criterion represent critical behaviors or outcomes of our position?

[4] According to Ford & Blair (2006), the boldfaced questions are especially critical to supporting the use of the test.
[5] According to Ford & Blair (2006), the boldfaced questions are especially critical to supporting the use of the test.

8. **What is the Standard Deviation Difference between relevant demographic groups (e.g., White-Black, White-Hispanic, Male-Female)?**
9. **Were steps taken to reduce bias in the collection of the criterion?**
10. What is the uncorrected validity coefficient(s)?
11. **If available, what is the corrected validity coefficient(s)?**
12. Was the validity coefficient corrected using appropriate methodology?

All of IPMA-HR's promotional tests are validated using *content-related* validation processes. As you read through our promotional technical reports, here are 12 questions to consider when evaluating content-related validation evidence (Ford & Blair, 2006): [6]

1. **Does the description of the job domain in the job analysis match the job domain for our position?**
2. **Did the job analysis include personal characteristics (e.g., other relevant characteristics such as interpersonal skills)?**
3. Does the job analysis report appropriately define KSAPs?
4. Are all the KSAPs required to perform critical job behaviors for our position included in the defined job domain?
5. **Does the test appropriately define the constructs it is measuring?**
6. Does the test measure most of the critical parts of the job domain?
7. **Is there any evidence demonstrating that the test accurately measures critical aspects of the job domain (i.e., the relevant KSAPs)?**
8. **Does the level of complexity for reading, math, and other skills/abilities approximate our job requirements?**
9. Does the test measure KSAPs that are learned on the job? KSAPs that are learned through training should not be assessed on an entry-level exam.

[6] According to Ford & Blair (2006), the boldfaced questions are especially critical to supporting the use of the test.

10. Are the time limits appropriate?
11. Are the weights appropriate for our position and agency (i.e., are they consistent with our job analysis or job description)?

As you read through our test response data reports for both entry-level and promotional tests, here are seven questions to consider when evaluating adverse impact evidence (Ford & Blair, 2006): [7]

1. What was the size of the sample (s) used to evaluate adverse impact?
2. **Is the adverse impact sample representative of our labor market?**
3. Was adverse impact evaluated in an incumbent or applicant sample?
4. What is the adverse impact for the test (e.g., white-black, white-Hispanic, male-female)?
5. **Is the pass point at which adverse impact is reported similar to our pass point?**
6. **Does the adverse impact ratio exceed the 4/5 rule? (i.e., 80%)?**

Determining the Relevant Labor Market: Who Is an Applicant?

Determination of the relevant labor market requires answering the following question: Who is an applicant? At first glance, answering this question may seem both simple and straightforward. An applicant is any individual who applies, right? However, it is not always clear when or how an individual becomes an applicant. Let's say an applicant is simply an individual who applies for the job. What does applying mean—talking to a recruiter, filling out a set of standardized forms, emailing a resume, meeting a predefined set of requirements? Does applying include some of these, all of these, none of these?

[7] According to Ford & Blair (2006), the boldfaced questions are especially critical to supporting the use of the test.

You may wonder why we are bothering to answer this question at all if it has such an ambiguous and monotonous answer. You may be further wondering why we are bothering to answer this question when we are supposed to be discussing adverse impact. Determining which individuals are applicants is a highly important issue in the context of selection and is directly related to our discussion of adverse impact. First, even if a test user has not collected data on adverse impact, federal enforcement agencies may infer the impact of a selection procedure based on the group's representation in the relevant labor market or the user's workforce (in the case of jobs filled internally).

Second, the definition of the term "applicant" is critical to several record-keeping provisions, as specified in the *Uniform Guidelines*.

> Under UGESP, employers 'should maintain and have available for inspection records or other information which will disclose the impact which its tests and other selection procedures have upon employment opportunities of persons by identifiable race, sex, or ethnic group. . . .' 29 C.F.R. 1607.4A. When hiring, each covered entity should maintain and have available such information with respect to applicants. When promoting, each covered entity should maintain and have available such information with respect to candidates. If tests and selection procedures are shown to have a disparate impact, then UGESP sets forth different methods for determining whether the tests or selection procedures are valid and job-related. UGESP is at 29 C.F.R. Part 1607 (EEOC) and 41. (taken 07/20/09 from http://www.ofccp.com/applicant.asp):

Third, making determinations about the relevant labor market is also important because it helps you understand your applicant pool. This can be important when you are making decisions regarding selection measures. For example, suppose you are reviewing the technical report for a test you are considering using in your upcoming selection process. If you know who your applicants are, you can compare the demographics of your agency's incumbents and labor market to the validation sample to be sure that the validation sample is a fair representation.

According to the Office of Federal Contract Compliance (OFFCP), there are four criteria used to determine who is an applicant (taken 07/20/09 from http://www.ofccp.com/applicant.asp):

Criteria 1 states that an individual is an applicant if he or she expresses interest in employment via the Internet or a similar technology. OFFCP notes several implications for employers who are federal contractors:

> Implication: Employers may want to include some electronic response option for all open positions since this new definition will allow all the applicants to be counted in the same way (and this new way is more favorable to employers than in the past, which required employers to count all those who responded to be counted as applicants, even though they did not meet basic qualifications). This means candidates who are not "basically qualified" will not be counted as applicants. Adding an email address as a response option permits employers to use all four eligibility requirements outlined in the new regulations, which is more favorable to employers.
>
> Implication: Individuals must express interest to the employer in a particular position in order to be a true applicant. Therefore, resumes generated by going to an on-line data base such as Monster would not be counted until specific resumes are culled through a search using the basic requirements of the job and those candidates have indicated mutual interest.

Criteria 2 states that an individual is an applicant if a contractor considers that individual for employment. OFFCP notes several implications for employers who are federal contractors:

Implication: This regulation is the same as the old regulation and specifies the employer use some sort of tool to screen candidates, whether manually reviewing qualifications or using electronic tools.

Criteria 3 states the individual is an applicant if his or her expression of interest indicates that he or she possesses the minimum qualifications to qualify for the position. OFFCP notes several implications for employers who are federal contractors:

Implication: Employers must use care in determining “basic qualifications” for the job, ensuring these are job-related, objective, and non- comparative (for example, not the top five candidates). Basic qualifications are substantive credentials established for the job, such as experience, education, certifications, and other objective credentials. Employers are obliged to list the basic requirements in any recruitment advertising and pre-determine these qualifications prior to the selection process. The OFCCP will closely examine an employer’s basic qualifications for those positions where there is adverse impact greater than two standard deviations.

Criteria 4 states the individual is an applicant if at no point in the contractor’s selection process prior to receiving an offer of employment from the contractor does that individual remove himself or herself from further consideration or otherwise indicate that he or she is no longer interested in the position. OFFCP notes several implications for employers who are federal contractors:

Implication: You don’t have to count individuals as applicants if they: reject an offer; fail to respond to inquiries; are unwilling to meet requirements of the job (e.g., shifts, overtime, travel); or indicate higher salary requirements than are afforded by this position.

Some employers are examining “front-loading” key screening questions early in the selection process, that is, before other credentials are “considered” by the employer. For example, if candidates are to supply a cover letter or complete a form indicating their minimum salary requirements and openness to various work

elements such as travel and overtime, then candidates who are not open to the specifications of the job can be eliminated from being counted as an applicant.

Implication: Employers may establish their own procedures for selection and may differentiate procedures for different types of positions. However, employers must remain consistent in applying their own procedures within these job categories.

Implication: You may limit the number of candidates you review on a non-comparative basis as a data management technique. For example, you may look at blocks of candidates (say, the first 20 who apply, or the last 20 who apply). While the regulations don't speak to the record-keeping requirements, it is recommended that employers keep a record of the technique used.

Ensuring Job-Relatedness

To be defensible in court, you must be able to demonstrate that a test is job-related. That is, you must be able to present evidence of its job-relatedness. If you are administering a test that was developed in-house, this can be accomplished by conducting a local validation study. If you are renting a test from a publisher, you can conduct a local validation study or consider conducting a transportability study. If the results of your local validation or transportability study do not show the test to be job-related, you will need to consider whether suitable alternatives exist. Please consult with your legal department for more details. If using a published test, review any adverse impact data available from the test publisher.

In many situations, it is not feasible to gather local validation evidence, due to practical or logistical constraints. There are several strategies that collect and generalize evidence of validity without conducting a full-scale validation study. The most appropriate method to demonstrate if a validation study conducted by IPMA-HR applies to your locality would be a transportability study. Transportability is a specific process whereby evidence of validity may be generalized from one situation to another, based on the results of a validation study conducted elsewhere. According to the *Principles*

(SIOP, 2003), in order to generalize evidence of validity elsewhere without conducting a local validation study, it is important to review the methodology used in the original validation study to ensure its technical soundness and to establish the relevance of the study to the situation at hand.

In doing so, be sure to examine the degree of similarity between the job performed by incumbents locally and the job performed at the location where the test has been used previously. This may be accomplished through the examination of task analysis data. To learn more about transportability studies, see Hoffman & McPhail (1998) and Hoffman (1999). Please note that IPMA-HR will offer its transportability study service again starting in early 2010.

Ensuring Adequate Test Preparation

Reading Lists. All of IPMA-HR's promotional tests are associated with reading lists. These lists contain textbooks or articles, which IPMA-HR strongly recommends to candidates to study prior to taking the exam. The material in the books bears direct relevance to the exam content. Candidates should receive the list of books and then have 60 to 90 days[8] to study them. Distribution of the reading list is important because (a) it forces candidates to learn or reinforce job knowledge and (b) it creates a non-litigious atmosphere where candidates are treated fairly and given the highest chance of success. For example, candidates may not have learned the material that is covered by these books on the job, and their only exposure to it would be during this study period. IPMA-HR has received distressed telephone calls and e-mails from agencies reporting high failure rates on promotional exams. When this happens, one of the first questions we ask is whether the reading list was distributed to candidates. More often than not, the answer is "no" and that the person calling had no idea the reading list even existed. The purpose of this warning is not to embarrass or criticize anyone, but rather to encourage awareness of the importance of reading lists.

[8] Please build adequate time into your testing schedule for candidates to both *order* and *study* the books.

Study Guides. Be sure to ask your IPMA-HR customer service representative whether there are study guides available for the test you are administering. Our study guides are designed to eliminate the veil of mystery that often surrounds test day. They are intended to decrease test anxiety by providing candidates with sample questions, and by familiarizing them with general instructions and question formats. Theoretically, familiarizing candidates with these aspects of the test allows for a more accurate assessment of the candidate's knowledge, skills, and abilities. That is, once candidates are familiar with the structure of the test, they are free to focus on the content itself. Since you are no longer assessing how well they take tests, you are left with a truer picture of their actual abilities.

Currently, the *P-Series Entry-Level Police Officer Candidate Study Guide* is available for purchase by agencies and candidates. Please request an informational handout, which describes details including how much they cost and where to buy them. These handouts can be distributed directly to candidates.

Test Information Packets. Some of IPMA-HR's entry-level tests are accompanied by Test Information Packets (TIPs), which candidates are instructed to study 20 to 25 minutes prior to starting the actual test. These packets contain information that candidates will be tested on later, once the actual test begins. Candidates are not permitted to take notes, but you should refer to the TIP itself for specific information. TIPs are designed so that the average reader has time to read the packet twice. In fact, the directions in each TIP clearly state that candidates should reread the material if they have time. If the test you are planning to administer has a TIP, you may want to reiterate to candidates before distributing the TIP that they should reread the material if time permits. Likewise, it is important to observe whether candidates appear to be taking this part of the exam seriously.

Good Old-Fashioned Encouragement. Consider providing candidates with some common-sense advice, such as encouraging them to read questions carefully, study directions thoroughly, use the full time limits, check their work if they have time, etc. If

for no other reason, this may give candidates a better experience overall, and they will have less cause for complaint later.

ADVERSE IMPACT ANALYSES

According to Biddle (2006), there are two primary types of adverse impact analyses—Selection Rate Comparisons and Availability Comparisons. A Selection Rate Comparison can be used to evaluate the selection rates between two groups—a focal group and a reference group. "Selection Rate Comparisons are most typically used in litigation settings, as they relate specifically to the type of adverse impact analysis called for in the *Uniform Guidelines*" (Biddle, 2006, p. 5). Focal group members usually encompass females or minorities, while reference group members usually include males or whites. Four types of variables are involved in Selection Rate Comparisons: the number of focal group members, the number of focal group members who were not selected, the number of reference group members who were selected, and the number of reference group members who were not selected. Please note that this type of analysis can also be used to compare two groups in other instances, such as layoffs, promotions, or placements (Biddle, 2006).

An Availability Comparison evaluates one group's representation in a position to their availability for that position. Again, focal group members usually encompass females or minorities, while reference group members usually include males or whites. For example, let us pretend that 10% of incumbents in the Police Lieutenant position are female, while 15% of qualified applicants available for Police Lieutenant are female. According to Biddle (2006), Availability Comparisons are useful for demonstrating the extent to which one group is utilized. In our example, 15% of qualified females are available and only 10% are being utilized, which indicates a 5% underutilization of females. Three types of variables are involved in Availability Comparisons: the total number of incumbents in the position, the number of focal group members in the position, and the percentage of qualified focal group members who are available for the position.

While both the Selection Rate Comparison and Availability Comparison can be used to gauge information about the constitution of your work force, how the results can be used and what the results mean in terms of consequences for your agency differ greatly. "There are major differences between these two types of analyses regarding the extent to which they constitute legitimate finding of adverse impact that could potentially bring an employer to court. Generally speaking, the Selection Rate Comparison is the only type that can be used alone to demonstrate adverse impact in the classical sense; whereas the Availability Comparison only shows a prima facie reason to investigate further into an employer's practices to see why a "gap" may exist" (Biddle, 2006, p. 5).

Perhaps, a practical and useful way to incorporate them into your selection system would be as follows. Periodically, compute Availability Comparisons. This may help inform your Affirmative Action Plan, if you have one. Following the administration of a selection procedure, make Selection Rate Comparisons. While it is illegal to make decisions about hiring and promotion based on demographics alone, it is perfectly acceptable to be aware of your Affirmative Action goals and whether you meet them. Following a test administration, for instance, if you compute Selection Rate Comparisons and Availability Comparisons and find that the results do not meet your Affirmative Action goals, the analyses can be used to help justify modifications you may want to make to your selection procedures before the next round of hiring or promotions begins.

Statistical Significance

Although both descriptive statistics and statistical significance tests can be applied to Selection Rate Comparisons and Availability Comparisons, statistical significance tests are more pertinent for adverse impact analyses. "Statistical significance tests are the most relevant for adverse impact analyses because they show whether the descriptive statistic is statistically meaningful and whether they can be regarded as a 'beyond chance' occurrence…Statistical significance tests result in a p-value (for probability). P-values range from 0 to +1. A p-value of 0.01 means that the odds of the event occurring

by chance is only 1%. A p-value of 1.0 means that there is essentially a 100% certainty that the event is 'merely a chance occurrence,' and cannot be considered as a 'meaningful finding.' P-values of .05 or less are said to be 'statistically significant' in the realm of EEO analyses. This .05 level (or 5%) corresponds with the odds ratio of '1 chance in 20'" (Biddle, 2006, p. 6).

Statistical Power

Statistical power is another important concept in adverse impact. Statistical power is the ability to reveal a statistical significant finding within a dataset, if there is one to be found. That is, a researcher's ability to find statistical significance largely depends on whether the analysis has enough power to reveal it. "Thus, a powerful adverse impact analysis is one that has a high likelihood of uncovering adverse impact if it really exists" (Biddle, 2006, p. 7). With regard to adverse impact analyses, there are three factors that affect the statistical power of the analysis: effect size, sample size, and the type of statistic test used (Biddle, 2006).

Effect Size. In the context of adverse impact analyses, the effect size for Selection Rate Comparisons is the numerical difference between the selection rates of the two groups. For example, if the selection rate for Hispanic males is 15% and the selection rate for white males is 70%, then the effect size would be 55%. The effect size for Availability Comparisons is the numerical difference between a group's representation in the workplace and that group's availability. Continuing with our previous example, if 15% of the workforce is composed of Hispanic males and 30% of Hispanic males are available, then the effect size is 15%.

Sample Size. Sample size refers to the number of observations. For example, if your focal group sample size is 12, then there are 12 individuals in the focal group. Sample size is directly related to statistical power. That is, the larger the sample size, the more statistical power there will be in your analysis and vice versa. It is important to note that samples that are too large or too small can yield inaccurate results and may have significant consequences in terms of litigation. "In litigation settings, it is almost always

more favorable to the employer to limit the sample size used in the adverse impact analysis (because this limits the statistical power and makes it more difficult to identify significant findings) and almost always more beneficial for the plaintiffs to increase the sample size (for the reverse reason). The courts will typically have the final say on the nature of the sample used…despite years of debate among statistical practitioners, there is no absolute, bottom-line threshold regarding the minimum sample size necessary for conducting statistical investigations. In particular, the courts frequently take the stance that there is no clear minimum sample size. For example, in *Bradley v. Pizzaco of Nebraska, Inc.* (1991), the court stated 'there is no minimum size prescribed either in federal law or statistical theory" (Biddle, 2006, pp. 8-9).

If you want to increase the statistical power of an adverse impact analysis, consider combining different strata within the dataset (Biddle, 2006). Several strategies exist for doing this. According to the *Uniform Guidelines*, for example, "where the user's evidence concerning the impact of a selection procedure indicates adverse impact but is based upon numbers which are too small to be reliable, evidence concerning the impact of the procedure over a longer period of time and/or evidence concerning the impact which the selection procedure had when used in the same manner in similar circumstances elsewhere may be considered in determining adverse impact" (EEOC, CSC, DOL, & DOJ, 1978, Section 4D). Sample size can also be increased by combining smaller datasets from a local test administration across geographic areas into a larger regional analysis by combining events from several jobs, job groups, or divisions; by combining the results of different selection procedures into an analysis of the overall selection process; or by combining different races/ethnic group into one larger group (Biddle, 2006). For more information, see Biddle (2006).

Although there is no statistically or legally mandated minimum threshold, the minimum sample size for adverse impact analysis is generally agreed to be between 20 and 30:

> Despite years of debate among statistical practitioners, there is no absolute bottom-line threshold regarding the minimum sample size necessary for conducting statistical investigations. In particular, the courts frequently take the

stance that there is no clear minimum sample size. For example, in *Bradley v. Pizzaco of Nebraska, Inc.* (1991), the court stated "There is no minimum sample size prescribed either in federal law or statistical theory."

If one had to pick a firm number for adverse impact analyses, it would be 30 with at least 5 expected for selection (i.e., hired, promoted, etc.) (OFCCP, 1993). The *Uniform Guidelines Questions & Answers* (#20) state that a sample of 20 is too small. In some circumstances, however, it can be argued that a sample with fewer than the "30 requirement" from OFCCP or the "20 requirement" from the *Uniform Guidelines* can still allow meaningful statistical analysis to be conducted. (Biddle, 2006, p. 9)

APPROACHES TO MINIMIZING ADVERSE IMPACT

Approaches to minimizing adverse impact with regard to the selection instrument used typically involve the consideration of cut scores, incorporating non-cognitive measures into a test battery, or using tests with non-traditional formats (e.g., video or audio-based tests). As a test user, it is extremely important that you provide the proper study materials to your candidates

Setting Cut Scores

Setting a cut score or passing point for the test is one of the most important decisions a test administrator can make. By setting a cut score, you are saying that candidates must have minimum level of proficiency on the test in order to progress in the selection or promotion process.

There are a number of factors to consider when setting a cut score and many ways to do it. For more on setting cut scores and techniques, see *Considerations in Implementing Selection Procedures*. Whichever technique you use, you must have defensible justification for setting it at a certain point. The consideration of cut scores is important in the current discussion, because setting them arbitrarily high can lead to adverse impact.

Incorporating Non-Cognitive Tests and Alternative Formats

Non-cognitive tests assess affective, personal, and social variables that predict intelligent job performance. Examples include interpersonal skills, emotional intelligence, cooperativeness, sense of responsibility, service-mindedness, and conscientiousness. Audio- and video-based tests present information in a different format than traditional paper-and-pencil tests and tend to reduce the extent to which reading comprehension is assessed. Many of IPMA-HR's entry-level tests incorporate these techniques. These include the P-BDQ Police Background Data Questionnaire, C-BDQ Corrections Background Data Questionnaire, A-4 Video-Based Police Officer Test, PST 2.0 Public Safety Telecommunicator Test, B-5 Firefighter Test, and all of the D-Series Tests.

Wiesen (2007) argues that current facial recognition items are typically line drawings that:

- Have nearly identical faces and expressions,
- Are black and white images,
- Have too few faces to make comparisons,
- Do not use long-term memory.

To improve these items, Wiesen (2007) suggests:

- Using photos rather than line drawings,
- Presenting more faces for comparison,
- Ensuring that photos have different facial expressions and are taken from different perspectives,
- Using long-term memory.

CHECKING YOUR TEST SENSE

The following is a brief quiz that assesses your knowledge of the material provided.

The answers can be found at the bottom of page 47.

1. Title VII of the Civil Rights Act is the most frequently cited statutory basis for employment discrimination challenges involving cut scores.

   A. True
   B. False

2. Which of the following statements is **TRUE** about adverse impact?

   A. There is adverse impact in any employment situation in which the selection or hiring rate for any race, sex, or ethnic group is less than 80% of the selection rate for the applicant group with the highest selection rate.
   B. There is adverse impact in any employment situation in which the selection or hiring rate for any race, sex, or ethnic group is close to but not quite 80% of the selection rate for the applicant group with the highest selection rate, if the difference in the selection rates is statistically significant (i.e., unlikely due to chance) and practically significant.
   C. Adverse impact is not a property of a test; rather, it is an analysis of the results of employment decisions that may be influenced, in part, by a test or any number of other factors (e.g., number of vacancies, seniority, performance ratings, budgetary constraints, etc.).
   D. All of the above.

3. Selection Rate Comparisons and Availability Comparisons are two main types of adverse impact analyses.

   A. True
   B. False

4. Which of the following are important concepts in adverse impact analyses?

   A. Sample size.

   B. Statistical power.

   C. Statistical significance.

   D. Effect size.

   E. All of the above.

5. Which of the following ways can assist in minimizing adverse impact?

   A. Setting cut scores at the appropriate level.

   B. Incorporating non-cognitive measures into your test battery.

   C. Using tests with non-traditional formats (e.g., video or audio-based tests).

   D. Providing proper study materials to your candidates.

   E. All of the above.

6 Match each phrase (see left) with the most appropriate definition (see right).

Write your answer in the space provided:

Phrase		**Definition**
6a___Reading List	i.	Provides information on the job analysis, test development, and validation process for a given test.
6b___Technical Report	ii.	Presents test data gathered on candidates from jurisdictions that use IPMA-HR tests.
6c___Test Information Packet (TIP)	iii.	Contains materials that will be assessed by the test and is distributed just prior to handing out the test booklets.
6d___Test Response Data Report	iv.	States text books used by candidates to prepare for promotional tests.

Answers

1. A 2. D 3. A 4. E 5. E 6a. iv 6b. i 6c. iii 6d. ii

REFERENCES

Adopted by the Equal Employment Opportunity Commission, the Civil Service Commission, the Department of Labor, and the Department of Justice (1978). Uniform guidelines on employee selection procedures. *Federal Register*, 43(166).

Adverse Impact home page. Retrieved December 5, 2008 from http://www.adverseimpact.org.

Albemarle Paper Co. v. Moody, 442 U.S. 405 (1975).

Biddle, D. (2006). *Adverse impact and test validation: A practitioner's guide to valid and defensible employment testing*. Burlington, VT: Gower Publishing Company.

Civil Rights Act of 1964, Title VII, 42 U.S.C. §2000*e et seq*. (2003).

Civil Rights Act of 1991 (CRA91). Pub. L. 102-166.

Connecticut v. Teal, 457 U.S. 440 (1982).

Dipboye, R.L. (2005). *Discrimination at Work: The psychological and organizational bases*. Mahwah, NJ: Lawrence Erlbaum Associates.

FindLaw for Legal Professionals. *Connecticut v. Teal*, 457 U.S. 440 (1982). Retrieved April 8, 2009 from FindLaw for Legal Professionals via http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=457&invol=440.

FindLaw for Legal Professionals. *Griggs v. Duke Power Co.,* 401 U.S. 424, 426, (1971). Retrieved February 9, 2009 from FindLaw for Legal Professionals via http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=401&invol=424.

FindLaw for Legal Professionals. *Washington (Mayor, DC) v. Davis*, 426 U.S. 229 (1976). Retrieved February 9, 2009 from FindLaw for Legal Professionals via http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=426&invol=229.

Ford, J.M. & Blair, M.D. (2006, June). *Approaches to addressing adverse impact: Opportunities, facades and pitfalls.* Breakout session presented at the meeting of the International Public Management Association Assessment Council, Las Vegas, NV.

Griggs v. Duke Power Co., 401 U.S. 424, 426, (1971).

Gutman A. (2005). *Adverse impact: Judicial, regulatory, and statutory authority*. In F.J. Landy's (Ed.), Employment Discrimination Litigation: Behavioral, Quantitative, and Legal Perspectives (pp. -20-46). San Francisco, CA: Jossey-Bass.

Kehoe, J.F., & Olson, A. (2005). Cut scores and employment discrimination litigation. In F.J. Landy's (Ed.) *Employment discrimination litigation: Behavioral, quantitative, and legal perspectives* (pp. 410-449). San Francisco: Jossey-Bass.

Landy, F.J. (2005). *Employment Discrimination Litigation*: *Behavioral, quantitative, and legal perspectives*. San Francisco, CA: Jossey-Bass.

Lawshe, C.H. (1987). Adverse impact: Is it a viable concept? Professional Psychology. Research and Practice, 18(5), 492-497.

Martin v. Wilks, 490 U.S. 755 (1989).

Mueller, L., Norris, D., and Oppler, S. (2005). Implementation Based on Alternative Validation Procedures: Ranking, Cut Scores, Banding, and Compensatory Models in S. Morton McPhail's (Ed.) Alternative Validation Strategies: Developing New and Leveraging Existing Validity Evidence. San Francisco: Jossey-Bass.

Our Documents (n.d.). Civil Rights Act (1964). Retrieved December 5, 2008. From http://www.ourdocuments.gov/doc.php?doc=97.

Patterson v. McLean Credit Union, 491 U.S. 164 (1989).

Price Waterhouse v. Hopkins, 490 U.S. 228 (1989).

Society for Industrial Organizational Psychology (2003). *Principles for the validation and use of personnel selection procedures, fourth edition*. Society for Industrial Organizational Psychology: Bowing Green, OH.

Wards Cove Packing Co. v. Atonio, 768 F.2d 1120 (1993).

Washington (Mayor, DC) v. Davis, 426 U.S. 229 (1976).