

Employee Selection Procedures and the Business Necessity Defense

Richard A. Posthuma
University of Texas at El Paso

This article describes the legal framework for the justification of employee selection procedures. It describes the concept of adverse impact and explains how employers must provide a business necessity for the use of selection procedures with adverse impact under the EEOC Uniform Guidelines on Employee Selection Procedures. Adopted in 1978, there have been several changes in the law of employee selection procedures. Some of these changes arose from judicial decision making. Others arose from legislative action. In light of these changes, this article was prepared to provide a concise and comprehensive summary of the current legal framework surrounding employee selection procedures.

Employee selection procedures are most often challenged under the provisions of Title VII of the Civil Rights Act of 1964 as amended. That statute has been interpreted to include two theories of discrimination (disparate treatment and disparate impact) each with its own defense (bona fide occupational qualification and business necessity). These theories and defenses are explained below.

Disparate Treatment

In 1964 the U.S. Congress passed a statute that was designed to ensure equal employment opportunities for persons regardless of their race, color, religion, sex, or national origin. This statute applies to employers, employment agencies, and labor unions and prevents discrimination in employment practices (hiring, firing, compensation, or other terms and conditions of employment). At first, the statute was applied to intentional acts of discrimination that later became known as disparate treatment cases. In this type of case there is evidence that the employer intended to treat persons differently because of their membership in a protected class (e.g., race, sex). For example in *Phillips v. Martin Marietta Corp.* (1971) the U.S. Supreme Court found that when an employer refused to accept applications from women with pre-school-age children, there was disparate treatment discrimination based on sex. Employers can defend against charges of disparate treatment in selection cases by showing that the selection criteria is a bona fide occupational qualification (BFOQ).

Disparate Impact

The U. S. Supreme Court also ruled that employers could be held liable for unlawful discrimination even if they did not intend to discriminate. In *Griggs*

v. Duke Power Co. (1971) the Supreme Court established the disparate impact theory of discrimination. In *Griggs*, the court ruled that employee selection criteria that disproportionately exclude subgroups of the population may not be used unless the employer can establish a business justification for them. *Griggs* enabled plaintiffs to prove discrimination without proving that the employer had the intent to discriminate as had been required in disparate treatment cases. However, *Griggs* also enabled employers to defend the use of selection criteria that had disparate impact by showing that the criteria were job related.

In *Connecticut v. Teal* (1982) the Supreme Court elaborated on the disparate impact theory by ruling that an employer may not justify a selection practice that has disparate impact by merely showing that the overall selection process did not have adverse impact. Rather, each element of a selection procedure can be challenged. Seven years later the Supreme Court ruled that plaintiffs need more than just a showing of some statistical disparity within an employer's workforce to show disparate impact. Rather, they must show that a specific identified employment practice has a disparate impact and prove how that practice is the cause of the statistically disparate outcome (*Antonio v. Wards Cove Packing Co.*, 1989). The implication of these two cases is that "bottom line" statistics are not enough. Plaintiffs cannot prove their disparate impact case by showing a bottom line statistical disparity in the employer's workforce. Likewise, employers may not justify a multi-stage selection process in which one step has disparate impact, by showing that the overall process has no "bottom line" disparate impact against minorities.

The Civil Rights Act of 1991

Two years after the *Wards Cove* decision, the United States Congress passed the Civil Rights Act of 1991 to amend the Civil Rights Act of 1964 in several respects. Figure 1 shows the current Title VII language including the 1991 amendments. These amendments established for the first time specific statutory language defining the disparate impact theory of discrimination and the business necessity defense. That language specifies that plaintiffs can prevail on a disparate impact theory only if an employer ...

"...uses a particular employment practice that causes a disparate impact on the basis of race, color, religion, sex, national origin and the respondent [employer] fails to demonstrate that the challenged practice *is job related for the position in question and consistent with business necessity...*". (42 U.S.C. § 2000e-2(k)(1)(A)(i)). [emphasis added].

This statutory language is generally consistent with the prior federal court rulings that followed the principles of disparate impact set forth in *Griggs v. Duke Power* (1971).

However, the new statute also incorporates the *Wards Cove* ruling that a specific practice be identified and be shown to have adverse impact. The revised statute states that plaintiffs ...

"...shall demonstrate that each particular challenged employment practice causes a disparate impact, except that if the complaining party can demonstrate to

the court that the elements of a respondent's decision making process are not capable of separation for analysis, the decision making process may be analyzed as one employment practice.” (42 U.S.C. § 2000e-2(k)(1)(B)(i)).

This language requires that the plaintiffs specify the particular selection practice they are challenging. It also requires that the plaintiffs show a causal linkage between the practice and a disparate result for on a subgroup of the pool of qualified applicants for the position.

Figure 1. Excerpts from the Civil Rights Act of 1964 as amended by the Civil Rights Act of 1991, (42 USCA §§ 1981, 2000e-2 (1994))

(k) Burden of proof in disparate impact cases

- (1) (A) An unlawful employment practice based on disparate impact is established under this subchapter only if-
 - (i) a complaining party demonstrates that a respondent uses a particular employment practice that causes a disparate impact on the basis of race, color, religion, sex, national origin and the respondent fails to demonstrate that the challenged practice is job related for the position in question and consistent with business necessity; or
 - (ii) the complaining party makes the demonstration described in subparagraph (C) with respect to an alternative employment practice and the respondent refuses to adopt such alternative employment practice.
 - (B)
 - (i) With respect to demonstrating that a particular employment practice causes a disparate impact as described in subparagraph (A)(i), the complaining party shall demonstrate that each particular challenged employment practice causes a disparate impact, except that if the complaining party can demonstrate to the court that the elements of respondent's decision making process are not capable of separation for analysis, the decision making process may be analyzed as one employment practice.
 - (ii) If the respondent demonstrates that a specific employment practice does not cause the disparate impact, the respondent shall not be required to demonstrate that such practice is required by business necessity.
 - (C) The demonstration referred to by subparagraph (A)(ii) shall be in accordance with the law as it existed on June 4, 1989, with respect to the concept of "alternative employment practice."
- (2) A demonstration that an employment practice is required by business necessity may not be used as a defense against a claim of intentional discrimination under this subchapter.
 - (3) Notwithstanding any other provision of this subchapter, a rule barring the employment of an individual who currently and knowingly uses or possesses a controlled substance, as defined in schedules I and II of section 102(6) of the Controlled Substances Act (21 U.S.C. 802(6)), other than the use or possession of a drug taken under the supervision of a licensed health care professional, or any other use or possession authorized by the Controlled Substances Act [21 U.S.C.A. 801 et seq.] or any other provision of Federal law, shall be considered an unlawful employment practice under this subchapter only if such rule is adopted or applied with an intent to discriminate because of race, color, religion, sex, or national origin.

(l) Prohibition of discriminatory use of test scores

It shall be an unlawful employment practice for a respondent, in connection with the selection or referral of applicants or candidates for employment or promotion, to adjust the scores of, use different cut off scores for, or otherwise alter the results of, employment related tests on the basis of race, color, religion, sex, or national origin.

(m) Impermissible consideration of race, color, religion, sex, or national origin in employment practices

Except as otherwise provided in this subchapter, an unlawful employment practice is established when the complaining party demonstrates that race, color, religion, sex, or national origin was a motivating factor for any employment practice, even though other factors also motivated the practice.

Further, the 1991 amendments overturned the part of the *Wards Cove* (1989) decision that made it more difficult for plaintiffs to show that a selection practice was merely a pretext for unlawful discrimination. Prior to *Wards Cove*, the courts held that even if the employer proved a business justification for the practice that had adverse impact, the plaintiffs could still show that the employer should have used another selection practice that would have less adverse impact. The implication of the plaintiff's argument was that the employer's choice of a selection practice with more adverse impact shows that their purported job-relatedness defense was merely a pretext for unlawful discrimination.

As a result of the statutory framework and the judicial decisions, it is now clear that employers must be prepared to justify the use of their selection procedures if they have adverse impact. Although Title VII does not state explicitly how this justification must occur, many courts have used the EEOC Uniform Guidelines on Employee Selection Procedures as a guide in determining whether employers have justified their selection procedure.

The Uniform Guidelines on Employee Selection Procedures

The U.S. Equal Employment Opportunity Commission (EEOC) promulgated the Uniform Guidelines on Employee Selection Procedures in 1978 (EEOC, 1978). Based in large part on expert input from and standards promulgated by the Society of Industrial for Industrial and Organizational Psychology (SIOP, 1987) and the American Psychological Association (APA, 1985), these standards have guided both selection professionals and the courts for nearly a quarter of a century. While not having the binding force of a statute or official regulation, they have been widely used and applied. A search of Lexis-Nexis electronic database reveals that these Guidelines have been influential in more than 200 federal court decisions.

Recently, a joint effort by the American Educational Research Association, the APA, and the National Council on Measurement in Education has resulted in the promulgation of a new set of standards for testing at school and in the workplace (AERA, APA, NCME, 1999). There is also a group working on a revision of the SIOP standards for employee selection testing. These efforts may eventually lead to changes in the EEOC Guidelines. As a result, scholars and practitioners are taking a fresh look at the EEOC Guidelines.

The following summary provides a synopsis that digests the Guidelines into an easily accessible and understandable format. This outline was prepared using a modified Delphi research technique. In this method, each sentence of the text of the Uniform Guidelines was analyzed and matched to a specific category or topic. Then, the categories and topics were summarized into the coherent framework that is shown in the outline below. Embedded in the outline are comments that refer to important court decisions and federal legislation that modify or limit the applicability of the Guidelines. The summary is intended to be comprehensive in that it touches on all of the substantive matters addressed by the Guidelines. Thus, it may be a useful tool and checklist for those interested in selection procedure validation and justification.

Outline Summary of the EEOC Uniform Guidelines on Employee Selection Procedures

The following outline summarizes the provisions of the Uniform Guidelines and highlights the provisions of the guidelines that have been affected by subsequent judicial decisions or legislation.

I. Precursors to Validation

A. General

Necessary Data Collection. Employers should record data for each sex, race, and ethnic group that represents more than 2% of the labor pool. These data should include numbers of applicants and the number selected. The employer should record statistics for predictor and criterion variables for each subgroup such as: means, ranges, standard deviations, and correlations between predictor and criterion variables.

Feasibility of Studies. The feasibility of validation and fairness studies depends, in part, on the number of applicants and persons selected and the ranges of scores on predictor and criterion measures. In some cases, employers are relieved of the duty to complete validation and fairness studies where they are not feasible. However, they must generally take steps to do so. This may involve collecting data over longer time periods.

B. Adverse Impact & Fairness

Adverse Impact Analysis. Employers should determine whether scores on a selection procedure (e.g., interview, pencil and paper test, etc.) are lower for one subgroup than another and whether this difference results in a subgroup being selected less often for hiring, promotion, or other employment decisions. Each subgroup should be compared to the subgroup that has the highest passing rate. If the passing rate is less than 80% of the group with the highest passing rate, the selection procedure should be validated. However, this is not a fixed rule or quota and other statistics may also be used to show adverse impact or the lack thereof.

Adverse Impact of Elements of Selection Procedures. The Guidelines suggest that if a combination of procedures does not result in adverse impact, individual elements do not need to be evaluated. However, the U.S. Supreme Court in *Connecticut v. Teal* specifically rejected the bottom line defense. In addition, the Civil Rights Act of 1991 provides new language that addresses how employers should approach the analysis of individual elements and total selection procedures.

Need for Test Validation. Where a selection procedure is shown to have adverse impact, the test user should obtain evidence of the validity of the test. In

general, validity refers to a statistically significant relationship (e.g., correlation coefficient, $p < .05$) between predictor scores and criterion measures. However, test validation may be done by criterion (predictive or concurrent), content, construct, or some other strategy that follows professionally accepted principles.

Need for Fairness Analysis. Where a selection procedure is shown to have adverse impact, it should also be analyzed for fairness. Fairness means that persons in one subgroup obtain lower scores on a predictor measure, and the differences in scores are not reflected in differences in criterion scores.

II. Validation

A. General

Standardization Preferred. Employers should use standardized selection procedures as opposed to informal procedures that are not scored.

Employer's EEO Posture. An employer's general posture towards equal employment opportunity and affirmative action will be considered in evaluating the appropriateness of selection measures.

Compliance with Professional Standards. The validation of selection practices is enhanced by the showing of compliance with professional standards (e.g., APA, SIOP). However, the mere reputation of the party doing the validation or general advertising claims of test publishers do not support the conclusion that a test is valid or appropriate for use.

Appropriate Timing of Studies. In general, selection practices should be validated before they are used for selection decisions. However, in some cases, data on validation may be collected simultaneously with use. But if adverse impact results, the employer is not relieved from liability.

Adequacy and Completeness of Reports. Reports of test validation and fairness should include the important elements specified in the guidelines.

B. Job Analysis

Quality and Completeness of Job Analysis. A professionally prepared job analysis must be performed which will result in a description of those aspects of the job that are frequent, important, critical, and prerequisites to successful job performance. These could be work

behaviors, work outcomes, knowledge, skills, abilities, or other characteristics of a worker, worker behavior, or outcomes of the work.

C. Validation Procedures

1. Validation Strategies

Appropriateness of Validation Method. The selection validation method should be appropriate for the behaviors or outcomes to be measured. For example, validation of knowledge, skill, or ability constructs is best done by construct validation strategy. Multiple methods may be used.

Sufficiency of Empirical Evidence of Criterion Validation. Criterion validation should show an empirical relationship between test scores and job performance. No minimum effect size is required but higher effect sizes are preferred. Also, statistical significance should be greater than $p < .05$.

Matching Content of Test to Job for Content Validation. A content validation strategy must show that the content of the job matches necessary, important, or critical elements of the job itself.

If the content of prior training and/or experience is a predictor, the content of the specific behaviors, outcomes, and knowledge, skills, or abilities should match those included in the job. This is true even if the prior training and experience doesn't match the job as a whole.

Careful Explication of Constructs for Construct Validation. Construct validation requires careful explication of constructs. It also requires a series of studies, including criterion-related studies, that show a relationship between a construct and successful job performance.

Appropriateness of Statistical Adjustments. Care should be taken in making statistical score adjustments to avoid problems with capitalization on chance or overestimation of effects. For example, weighting, adjustments for low reliability or restriction in range, etc. may tend to inflate reported validities.

Appropriateness of Generalizability. Data from test publishers, other employers, or other jobs inside or outside the employer's workplace may be used where appropriate for validation or fairness analysis. The appropriateness of using this data depends on the similarity of the jobs and the context of the jobs studied to the jobs and context in which the tests will be used.

2. Issues Related to Criterion Measures.

Adequacy of Criterion Measures. Criterion measures may be scores on job behaviors, work outcomes, supervisor evaluations of employee behavior, or other measures that are shown to be important outcomes for an employer. However, the criterion measures should generally reflect frequent, necessary, important and/or critical aspects that reflect most of the job.

Avoidance of Bias in Criterion or Predictor Measures. Test validation should avoid possible biases in the collection of data on predictor and/or criterion measures. This should include avoiding the possibility that scores on predictor measures will influence rankings of employees on criterion measures.

Justification for Testing Things to be Learned on the Job. In general, tests should not be used to measure knowledge, skills, abilities, or other characteristics that applicants acquire during short periods after being hired.

Justification for Success in Training as a Criterion. Where success in a training program is a criterion for test validation, there should be a showing that the training program is linked to successful performance of the work itself (e.g., content or criterion-related evidence linking training success to job success).

Justification for Selection for Progressively Higher Level Positions. If part of the criterion will be successful performance in higher level positions than the position into which one is hired, there should be evidence of the frequency and probability of promotions to these higher level positions, job analysis of the higher level positions, and an analysis of whether the requirements of successful performance in the higher level positions may change in the future.

Adequacy of Explication of Constructs. If constructs are to be used in a validation strategy they should be carefully defined and distinguished from other constructs. This involves comparing and contrasting the identified constructs with other constructs and statistical techniques such as factor analysis.

Representativeness of Predictor and Criterion Samples. The data used to develop predictors and criterion should be obtained from persons of the same sex, race, and ethnic composition as the persons to whom the tests will be given.

3. Methods of Use and Application.

Appropriate Use of Test for Cut-Off, Ranking, or Banding.

The use of selection test scores for setting cut-off scores, ranking, or banding (grouping) individuals should be based on data that links the cut score, rank, or band to actual differences in job performance.

Retest Opportunity Provided. Applicants should be given opportunities to retest. However, test items should remain secure.

Appropriate Difficulty Level or Complexity and Contextual or Complicating Factors. Careful consideration should be given to the level of difficulty or complexity of test elements. There should be a relationship between the difficulty level or complexity of test elements and the difficulty level or complexity of expected job performance. Similarly, any factors in the context or manner in which the test is administered should be evaluated to ensure that they are similar to the context of performance.

Appropriateness of Techniques Used to Reduce Adverse Impact. Techniques used to reduce adverse impact (e.g., differential weighting) should be reported. However, the Civil Rights Act of 1991 prohibits adjusting the scores, using different cut off scores, or otherwise altering the results of employment related tests on the basis of race, color, religion, sex, or national origin.

Disparate Treatment Prohibited. Regardless of whether a test is valid, it may not be administered in a way that treats individuals differently based on their race, sex, or ethnicity. For example, you may not give a physical ability test only to women applicants when men apply for the same job.

III. Examination of Alternatives

A. Options to Use Alternative Predictors

Predictor tests may include scores on a job interview, pencil and paper selection test, work sample, or any other selection procedure or combination of procedures that is used to evaluate applicants.

B. Analysis of Alternatives

Where a selection procedure has adverse impact the test user should investigate alternative tests to determine if they have equal or greater utility and validity, but less adverse impact. In general, larger effect size and lower adverse impact is preferred. The Civil Rights Act of

1991 now mandates the consideration of test alternatives in certain circumstances.

Summary

This article describes the legal framework for the justification of employee selection procedures that have adverse impact. It also summarizes the key principles of the *Uniform Guidelines on Employee Selection* procedures which, if followed properly, can justify the use of selection procedures with adverse impact. In general, the Guidelines require employers to collect data to determine if their selection practices have adverse impact. If there is adverse impact, the Guidelines state that selection practices should be both validated and evaluated for fairness to the extent feasible. The Guidelines also specify how test validation should proceed. Finally, the Guidelines indicate that test users should consider the usefulness of alternative tests which may reduce adverse impact.

Several parts of the Guidelines have been superceded by provisions of The Civil Rights Act of 1991. The language of that statute now provides mandatory guidance to employers on employment test validation that supercede provisions included in the Guidelines.

The efforts of professional organizations in updating their testing guidelines may prompt the EEOC to consider a revision of their Guidelines as well. Perhaps by the 25th anniversary of the promulgation of the Guidelines there will be revisions that incorporate both the updated guidance of professional organizations such as the APA, SIOP, and the changes mandated by federal legislation and the courts.

References

American Psychological Association. (1985). *Standards for educational and psychological tests*. Washington, DC: Author.

American Psychological Association, Division 14. (1987). *Principles for validation and use of personnel selection procedures* (3rd ed.). Washington, DC: Author.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, D.C: Author.

Appelton v. Delotte & Touche LLP, 168 F.R.D. 221 (1996, M.D. TENN).

Civil Rights Act of 1991, Pub. L. No. 102-166, 105 Stat. 1071 (1991)

Connecticut v. Teal, 455 U.S. 903 (1982).

Equal Employment Opportunity Commission, U.S. Civil Service Commission, U.S. Department of Labor, & U.S. Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, 43 (166), 38290-38315.

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

Housey v. Carini Lincoln-Mercury, 817 F. Supp 762 (1993, E.D. WI).

Phillips v. Martin Marietta Corp. 400 U.S. 542 (1971).

Watson v. Fort Worth Bank & Trust, 487 U.S. 977 (1988).
Wards Cove Packing Co. v. Antonio, 490 U.S. 642. (1989).

Author Notes

1. Comments or questions about this article should be addressed to:
Richard A. Posthuma, Ph.D., J.D.
University of Texas at El Paso
College of Business Administration
El Paso, Texas 79912
(915) 747-8646
rposthuma@utep.edu

