

PERFORMANCE APPRAISAL AND TERMINATIONS: A REVIEW OF COURT DECISIONS SINCE BRITO V. ZIA WITH IMPLICATIONS FOR PERSONNEL PRACTICES

GERALD V. BARRETT
The University of Akron

MARY C. KERNAN
Kent State University

Court cases since the classic *Brito v. Zia* (1973) decision dealing with terminations based on subjective performance appraisals are reviewed. Professional interpretations of *Brito v. Zia* are also examined and criticized in light of professional practice and subsequent court decisions. Major themes and issues are distilled from the review of cases, and implications and recommendations for personnel practices were discussed.

Legal issues pertaining to personnel decisions have received a great deal of attention in recent years (Cascio & Bernardin, 1981; Faley, Kleiman & Lengnick-Hall, 1984; Feild & Holley, 1982; Holley & Feild, 1975; Kleiman & Durham, 1981). As a result of Title VII of the 1964 Civil Rights Act, any personnel practice that adversely affects certain classes of individuals is deemed unlawful unless justified by business necessity. The termination of an employee is probably the most critical personnel decision an organization can make. The United States is unique among industrialized countries in that terminations, or firings, are quite commonplace. While there are no accurate figures, approximately three million people are fired each year in the private sector (Stieber, 1984). Schreiber (1983) has estimated that 200,000 of these employees are fired "unfairly."

In addition to Title VII, other statutes have been enacted to protect employees against unjust termination and discrimination (Age Discrimination in Employment Act of 1967; National Labor Relations Act of 1935). In countering a charge of discrimination, the organization invariably claims that an alleged unfair termination resulted solely from poor job performance. Thus, a formal or informal performance appraisal is often the main defense an organization has against such a complaint.

The legal requirements for performance appraisals, and in particular, those used as a source of information upon which to base termination decisions, are less clear cut than requirements for other personnel practices

Correspondence and requests for reprints should be addressed to Gerald V. Barrett, Department of Psychology, The University of Akron, Akron, OH 44325.

such as initial selection. Although the *Uniform Guidelines on Employee Selection Procedures* (1978) apply to all procedures used for personnel decisions, they focus mainly on technical standards necessary for the validation of initial selection decisions (Kleiman & Durham, 1981). Performance appraisal is mentioned only in regard to its use as a criterion in empirical validation studies (Kleiman & Durham, 1981). Thus, the purpose of this paper is to review recent court decisions that deal with terminations based, at least in part, upon performance appraisals. Guidelines established for Title VII cases, standards advocated in professional journals and texts, and the use of subjective performance appraisal criteria are discussed. Judicial standards are compared with standards advocated in professional journals, with the nexus between the two constituting a major portion of this paper. Major themes and issues are distilled from a review of pertinent cases, and implications for personnel practices are discussed. It should be noted that performance appraisals used to make promotion decisions are not considered in this review. Cases dealing with this issue are discussed elsewhere and may have to meet different legal requirements than appraisals used for terminations (see Kleiman & Durham, 1981).

Brito v. Zia Co.

The most stringent legal standard applied to performance appraisal procedures has been the belief that they should be considered "tests" and therefore fall under the *Uniform Guidelines on Employee Selection Procedures* (1978). This standard arose out of *Brito v. Zia Co.* (1973), in which the court ruled that the *Guidelines* were applicable and must be followed in evaluating the adequacy of a performance appraisal instrument. In this case, Spanish-speaking subordinates were terminated as a result of their supervisors' performance appraisals. Supervisors rated subordinates on several work dimensions, including volume of work, quality of work, job knowledge, dependability, and cooperation. The court was critical of Zia's subjective appraisal process, noting that some supervisors doing the rating had never directly observed the work of employees and that raters lacked a single standard to apply to the dimensions of quality and quantity of work. The Zia Co. failed to present evidence concerning the validity of their performance evaluation "test," and the court was of the opinion that they should have presented *empirical* data demonstrating that the "test" was significantly correlated with actual work behavior.

It seems apparent that the court confused the concepts that are traditionally called tests and criteria in the field of personnel (Bass & Barrett, 1981). A performance appraisal instrument is traditionally considered to be not a test, but a criterion that is correlated with a test. In most organizations, the performance evaluation is the *only* standard by which an

employee's performance can be assessed. Therefore, it is nearly impossible to correlate this standard with any other external measure.

Because of its important implications for organizational practices, the Brito decision has attracted a good deal of attention among personnel specialists. In the following section, we discuss how some commentators have interpreted this decision and note several problems associated with these interpretations.

Professional Interpretations of Brito v. Zia

Several experts in the field of human resource management have taken the lead of *Brito v. Zia* (1973) and view the performance appraisal instrument as a test as defined under the *Uniform Guidelines* (1978) (Bernardin, Beatty, & Jensen, 1980; Cascio, 1982; Cascio & Bernardin, 1981; Feild & Holley, 1982). As such, they call for empirical validation of performance evaluation ratings. Cascio and Bernardin (1981) also maintain that each rater must be validated in addition to the appraisal system itself. Although Cascio and Bernardin state that it is extremely difficult to empirically demonstrate the validity of performance evaluation ratings, they propose a technique for validating raters where a rater would make a judgment on some known objective fact such as absenteeism and use that as evidence that the other dimensions, for which there were no empirical data, were also valid.

A second technique advocated is to correlate actual observations of ratee behavior with performance appraisal ratings. A drawback of both approaches is that they are extremely time consuming. Most important, however, these approaches would not prevent raters from discriminating on an individual basis. In other words, one could demonstrate the validity of a rater's performance on an overall basis by these techniques, but this would have no particular relevance if a rater wanted to discriminate in his or her ratings on *one* or more occasions. Moreover, one might convincingly argue that if an organization has objective data available (i.e., absenteeism records, actual instances of ratee behavior), they should use these data, rather than subjective ratings, for evaluations. The attempt to treat the criterion or the performance evaluation rating as a test negates the distinction between the two concepts. As early as 1950, Brogden and Taylor pointed out that the criterion (performance appraisal) could not be subjected to a satisfactory empirical test of its adequacy.

Others in the field maintain that a performance appraisal must be a content valid test supported by an empirical study (Holley & Feild, 1975; Latham & Wexley, 1981). This position has two main problems. First, according to the *Uniform Guidelines* (1978) and professional practice, the validity of a content valid test does not depend upon an empirical validity

study. In fact, several courts have accepted content valid tests that have not been supported by empirical validation (*Contreras v. City of Los Angeles*, 1982; *NEA v. South Carolina*, 1978).

Second, there are several critical differences between a content valid test and a performance appraisal instrument. One important difference is that an effort is usually made to construct an objective scoring system for a traditional paper-and-pencil content valid test. All candidates take the same test under the same standardized conditions with the score being determined in an objective fashion. In contrast, performance appraisals are not objectively scored. While an overall score can be obtained in each case, the underlying process in determining that number is quite different from that of a test. Since a performance evaluation may be performed by different supervisors, the score an individual receives represents an interaction between the individual's performance, the nature of the job, and the supervisor doing the actual rating.

Secondly, all items on a typical paper-and-pencil content valid test are given equal weight. Thus, a content valid test operates implicitly on a compensatory model. If an individual misses a number of items, he or she can still obtain a satisfactory score. On the other hand, the performance appraisal technique is implicitly based on a noncompensatory model. If an individual is rated low on one very important job dimension, receiving high ratings on other dimensions may not offset the low rating. Depending upon the nature of the job, low ratings on one dimension may be sufficient justification for termination.

To further illustrate the differences between a performance evaluation instrument and a test, consider a hypothetical performance appraisal form for the job of a nuclear power plant operator, which has 10 different dimensions. One dimension might be the proper operation of the reactor safety equipment. Perhaps all 30 operators have performed that job dimension in a safe fashion whenever they have been observed by their supervisor. This dimension, then, would have zero variance and would not correlate with any measure of objective performance or any of the remaining nine performance dimensions. However, the dimension is still extremely important for any termination decision. One would not want to argue that because the performance appraisal instrument does not have the usual psychometric properties of a test, the dimension of safe operation should not be used for evaluating employees.

Finally, since an empirical validation study conducted for a performance appraisal instrument would depend upon specific rater responses, whenever there is a new supervisor, a new validation study would be required. In contrast, a content valid test does not require continued reevaluation unless the nature of the job changes.

To summarize, the court's interpretation of the *Guidelines* in *Brito v. Zia* (1973) and their stress upon empirical validation has been accepted as the legal standard by many in the field of human resource management. While the *Guidelines* do warn that performance evaluations may be adversely influenced by supervisory biases, empirical validation is not required. In addition, the *Guidelines* and the courts do not demand empirical verification of content valid tests (*Contreras v. City of Los Angeles*, 1982; *NEA v. South Carolina*, 1978).

If the standard set forth in *Brito v. Zia* (1973) were applied to all subjective performance evaluations, then most would fail to meet it. For most jobs, there is simply no objective data that are available to "validate" the performance evaluation. Fortunately, as will be shown shortly, subsequent court decisions since *Brito v. Zia* (1973) have not required empirical validation of performance appraisals used for termination decisions.

The following section of the paper provides a review of court decisions since the classic *Brito* decision that concern terminations based on performance appraisals. An exhaustive search of LEXIS identified over 100 relevant cases. Fifty-one were selected for a more careful review since these cases dealt almost exclusively with subjective performance appraisals. Of these 51 cases, all but 10 were decided in favor of the organization. In discussing these decisions we focus on major themes and issues that the courts appear to consider in cases alleging discriminatory termination practices.

Decisions After Brito v. Zia

Though rarely cited, the *Brito v. Zia* (1973) case was reheard by the District Court in 1976 (*EEOC v. Zia Co. and Los Alamos Constructors*). The court ruled that Spanish-speaking Americans were not discriminated against with respect to terminations. As a direct result of the *Brito v. Zia* (1973) decision, the court heard a great deal of testimony about a newly developed performance appraisal process that had been devised by a consulting group and was subjective in part. The court was of the opinion that, although *Brito v. Zia* (1973) was critical of subjective standards, it did not demand the impossible. The court believed the defendants made a reasonable effort to develop an objective procedure and found the "testing" procedures (performance evaluations) adequate. This decision was later upheld by the appeals court in *EEOC v. Zia Co. and Los Alamos Constructors* (1978). Interestingly, there was no discussion of the *Uniform Guidelines* or empirical validity in either decision.

A recent United States Supreme Court decision in *Texas Department of Community Affairs v. Burdine* (1981) upheld the termination of a female employee who was discharged along with two of her co-workers because

they did not work well together. This case is widely recognized for reinforcing the four criteria outlined in *McDonnell Douglas v. Green* (1973) needed to establish a prima facie case of discrimination under the disparate treatment doctrine. Specifically, the plaintiff must show (1) that he/she belongs to a racial minority, (2) that he/she applied and was qualified for a job in which the employer was seeking applicants, (3) that despite his/her qualifications, he/she was rejected, and (4) that, after his/her rejection, the position remained open and the employer continued to seek applicants from persons of complainants' qualifications. However, the facts of the case also show that the decision reached by the organization was very subjective in nature in that Burdine's supervisor relied upon the advice of advisors and subordinates without a formal appraisal or procedure. In *Gupta v. IBM* (1980) the court also relied on the testimony of the plaintiff's supervisors and co-workers in deciding the case of a discharged Indian employee. Those who testified spoke of incidents that outlined Gupta's general incompetence as a systems engineer. No formal appraisal system was introduced concerning the plaintiff's on-the-job performance. The court ruled that Gupta was not a victim of discrimination but was justifiably discharged for failure to perform his job duties.

In 10 of the 51 decisions, the court ruled in favor of the plaintiff. Two central issues seem to emerge as determining factors in eight of these cases. In two cases (*Foster v. MCI*, 1983; *Johnson v. Olin Corporation*, 1980) the supervisors who conducted the performance appraisals were identified through court testimony as blatantly racist. For example, in *Johnson v. Olin Corporation* (1980) the court felt that the evaluation of the plaintiff's work performance by her supervisor was merely a pretext to effectuate his discriminatory attitudes. Credible evidence was presented regarding the supervisor's racial antagonism against blacks. Due to the organization's subjective appraisal system and a lack of review by upper management, the supervisor was free to act out his prejudices.

A second factor that led to unfavorable decisions for the organization concerns the uneven application of performance appraisal standards to minority and majority employees. In *Spiva v. Copperweld* (1980) the plaintiff was terminated for violating a "last-chance agreement." Determinations of violations were not based on any objective standards but were left entirely to the discretion of supervisory personnel. The plaintiff was able to demonstrate that white employees were treated more favorably with respect to last-chance agreements regarding absences and instances of tardiness and were retained by Copperweld even though their violations were more frequent and severe.

In *Bledsoe v. Wilker Bros.* (1980) the employer claimed that the plaintiff was terminated primarily for her inadequate typing. However, typing samples of majority group members who were retained were judged to be

incompetent, inadequate, or marginal by the defendant's own standards. Thus, standards for performance were not applied equally to blacks and whites. In a similar vein, discretion was unevenly exercised when rating the performance of black and white employees in *Lilly v. Harris Teeter Supermarkets* (1983), and the appeals court upheld the district court's finding of racial discrimination.

Three cases in which the courts ruled against the organization dealt with reverse discrimination (*Lanphear v. Prokop*, 1983; *Planells v. Howard University*, 1983; *Turgeon v. Howard University*, 1983). In *Planells v. Howard University* (1983) and *Turgeon v. Howard University* (1983) the plaintiffs were employed as assistant professors and were subsequently denied reappointment. Interestingly, much of the evidence offered by the university clearly supported the plaintiffs. In terms of publication criteria and teaching evaluations, the plaintiffs were considered as good or better than several black faculty members whose contracts were renewed. In the *Turgeon* case, the department also failed to properly consider her for reappointment as required by its own written policies and procedures. The court ruled that Howard University applied its standards for reappointment more stringently to the plaintiffs than to black faculty members.

In order to circumvent litigation, some personnel specialists view the use of objective appraisal criteria as a safer strategy than the use of subjective standards. However, the ninth case decided in favor of the plaintiff (*Weahkee v. Perry*, 1978) illustrates a common problem involved in formulating objective performance evaluations. Ostensibly objective criteria can often become subjective in nature. In *Weahkee v. Perry* (1978), the plaintiff was an American Indian and an employee of the EEOC. After two years with the agency, his job was converted and a production requirement of four cases per month was established as an "objective" performance standard. Since the plaintiff was a senior investigator, he was assigned the more difficult cases and subsequently failed to meet the newly established standard. It is evident that the organization's objective criterion must ultimately be weighed subjectively because cases are not equivalent in terms of amount of work required. Three months after receiving the unsatisfactory performance rating, the plaintiff was discharged.

The Equal Employment Opportunity Commission's investigation of discrimination in its own agency found that the production policy permitted a distortion in the production records, and therefore it was unfair to compare different investigators. The court upheld a favorable decision from the Federal Employee Appeals Authority, and the plaintiff was reinstated into the Equal Employment Opportunity Commission's work force. This case illustrates that even the Equal Employment Opportunity Commission can experience a difficult time with performance evaluations. Their attempt to devise a so-called objective performance appraisal system was found to be

faulty by their own internal investigation. Somewhat ironically, there was no attempt to validate their process empirically, nor did they reference the *Guidelines*.

Finally, in *Nord v. U.S. Steel* (1985) the plaintiff successfully demonstrated that she had a record of good performance appraisals before she requested a promotion. After this request her appraisals became negative, which led to her eventual termination. The court reasoned that the downgraded evaluations after her promotion request were evidence of discrimination.

Employers in the remaining 37 cases successfully defended their termination decisions. Aside from the importance of applying performance appraisal standards equally to all employees, three major themes were identified from a review of these cases:

1. Courts refuse to act as super personnel departments. They do not wish to decide if performance appraisal systems are just or unfair or whether they conform in all respects to accepted professional standards. They only wish to rule on whether or not performance appraisals *discriminate*. That is, does the organization treat similarly situated individuals differently.
2. Courts seem to look favorably on the use of a review system by upper-level personnel to prevent individual bias.
3. Courts also react favorably to the use of performance counseling designed to help employees improve substandard performance.

Plaintiffs often challenge subjective performance appraisals purely on the grounds that they are *per se* discriminatory since they involve supervisory judgments. The first theme addresses this issue. At one extreme, some courts seem to accept defendants' obviously flawed subjective performance evaluations. The appeals court in *Jones v. The Pitt County Board of Education* (1975) upheld the district court's finding that there was no racial motivation in the termination of a black school teacher. The plaintiff had taught for over 10 years in the school system, and when moved to a new location, she was evaluated by her principal. The principal's overall conclusion was that her performance was unsatisfactory. This conclusion was particularly interesting since, as the dissenting judge pointed out, she had been evaluated on four different occasions on a Teacher's Evaluation Form. Out of a total of 165 ratings, she received only 5 weak scores. The principal's evaluation, which was supposed to be based on the Teacher's Evaluation Form, found her weak in 11 of 33 categories. This would appear to be a case where an expert witness was needed to point out the deficiencies in the principal's evaluation of the teacher.

However, even if ratings are open to divergent opinions and are challenged by expert testimony, courts often hold that ratings are not unduly subjective or arbitrary (*Harkless v. Sweeny School District*, 1975; *Marshall v. Kirkland*, 1979). In *Harkless v. Sweeny School District* (1975) the

court accepted a faculty evaluation form, which was essentially a rating of 38 characteristics on a scale of excellent, good, above the average, fair or average, poor, below the average, and unsatisfactory. The plaintiff's expert witness maintained that the evaluation technique was invalid. All requirements that the expert thought to be important for the performance evaluation process were those that were deficient in the one used by the defendant (i.e., rater training, multiple raters). The court was of the opinion that the evaluation procedures were not unfair, biased, or fatally subjective even though they did not conform in all respects to the expert's standards and stated that "The question before the court is not whether the District practiced perfect or even good techniques of managerial decision making. It is, rather, whether the District practiced racial discrimination in such decision making." The court also rejected the use of more objective "paper" qualifications, such as level of state certification and years of teaching experience. Similarly, in *Walston v. County School Board of Nansemond County* (1974) the court preferred a subjective performance evaluation over more objective criteria.

In *Kresyman v. Bolger* (1984), the plaintiff was terminated at the end of a 90-day probationary period for being slow on special assignments, being insensitive to co-workers and having two unscheduled absences. Testimony at the trial indicated that the incidents leading to the three unfavorable evaluations were very minor in nature. The court noted that "if it had before it the question solely of whether or not defendant had 'just cause' to terminate plaintiff, the court's finding might be for the plaintiff. But this is not the relevant inquiry. The court must determine whether there has been racial discrimination. The answer is no." Two other district court decisions (*EEOC v. Murphy Motor Freight*, 1980; *Evans v. Meadow Steel Products, Inc.*, 1984) and four appellate level decisions (*Mason v. Pierce*, 1985; *Nath v. General Electric*, 1977; *Powell v. Syracuse University*, 1978; *Powers v. Mancos School District RE-6 Montezuma County*, 1976) adopted similar lines of reasoning in finding that the respective plaintiffs were not victims of discrimination.

The second theme, review of performance appraisals by upper-level personnel, emerged as an important factor in several decisions. In *Richards v. IBM* (1981), the court was impressed that prior to the plaintiff's termination, his performance appraisals were discussed and reviewed by the supervisor's manager, a third-level manager, and the regional EEO manager. In *Grant v. C&P Telephone Co.* (1984), the plaintiff's work records were not only reviewed by upper-level personnel, he was warned repeatedly that termination was imminent if his work did not improve.

In *James v. Commonwealth Dept. of Public Welfare* (1983), the decision to discharge two black females for their poor job performance shortly after they completed an on-the-job training program was scrutinized as part

of a four-stage review process. Review began with the training supervisor and continued with the line supervisor, and there was an evaluation by the training director. The final termination decision was made by an independent executive board. This review procedure was considered effective in preventing intentional bias by any one individual. Review through higher levels of management and prior warning of termination were also important in *Carter v. Dialysis Clinic* (1981) and *Sias v. G.E. Information Services Co.* (1981).

Finally, several organizations who successfully defended their termination decisions offered performance counseling or corrective guidance to poor performers prior to reaching the final discharge decision.

In *Stone v. Xerox* (1982) the organization had a fairly elaborate procedure for assisting poor performers. The plaintiff was employed as a sales representative and in less than six months had been given several written reprimands concerning customer complaints about his selling methods and failure to develop adequate written selling proposals. As a result, he was placed on a one-month performance improvement program designed to correct these deficiencies. This program was extended 30 days at the plaintiff's request. Performance did not improve, and the plaintiff was then placed on probation and specifically told that failure to substantially improve would result in termination. The plaintiff's performance continued to be substandard, and he was discharged at the end of the probationary period.

Some form of performance counseling designed to assist employees in improving poor performance was also employed by organizations in *Bailey v. MCI Telecommunications Corp.* (1982), *Bennett v. Eggers* (1981), *Blunt v. Marion County School Board* (1975), *Bradington v. IBM* (1974), *Buchanan v. American Petroleum Institute* (1981), *Gidion v. Hospital for Joint Diseases* (1979), *Harris v. Group Health Association* (1981), *Howard v. Miller Brewing Co.* (1983), and *Sias v. G.E. Information Services Co.* (1981). For example, in *Howard v. Miller Brewing Co.* (1983), the plaintiff was counseled by a supervisor and advised on how to improve his performance after being rated "marginal in all respects." Subsequent ratings six months later indicated no improvement, and the plaintiff was terminated. In two cases (*Blunt v. Marion County School Board*, 1975; *Buchanan v. American Petroleum Institute*, 1981) corrective guidance was offered by the employer and was ignored, resulting in termination.

Conclusions and Implications for Personnel Practices

A review of pertinent case law can help identify what aspects of performance appraisals the courts routinely consider in deciding cases alleging discriminatory termination practices. In reviewing the relevant cases, there

does not appear to be a great deal of consistency in how the various courts determine the adequacy of a performance appraisal procedure. Certainly, much of the argument centers on plaintiffs challenging the use of subjective standards and the concomitant belief that supervisory judgments foster unfair, biased, and discriminatory appraisals. It seems clear, however, that the courts do not reject the subjective approach. In fact, some courts openly advocate subjective appraisals of employee attitudes and work behaviors. This is advantageous from an organization's point of view since most performance appraisals, especially for higher-level jobs, must ultimately rest on supervisory subjectivity. As seen in *Weahkee v. Perry* (1978) the use of objective performance appraisal criteria does not necessarily prevent litigation or ensure a favorable decision for the organization.

Decisions after *Brito v. Zia* (1973) do not view the performance appraisal as a "test" in cases alleging discriminatory terminations. The courts have not referred to the *Uniform Guidelines* as a standard which should be followed in evaluating a performance appraisal system in Title VII cases, nor do they require that a performance appraisal be considered a content valid test or that there be an independent empirical verification of the performance appraisal process. In sharp contrast, many commentators have followed the lead of *Brito v. Zia* (1973) by advocating a content validity strategy or empirical validation of performance appraisal procedures.

The professional literature concerning performance appraisals advocates and discusses extremely technical approaches to this process. These include refined techniques to develop a better performance appraisal form, reliability, training of raters to reduce rating errors, policy capturing, and validation techniques (Bernardin & Buckley, 1981; Cornelius, Hakel, & Sackett, 1979; Hobson, Mendel, & Gibson, 1981; Landy & Farr, 1980; Latham, Fay, & Saari, 1979; Taylor & Wilstead, 1974). Ironically, none of the court cases went into this sort of depth in determining the adequacy of performance appraisals. It appears evident that the courts are not often impressed by testimony that deals with the technical and psychometric properties of performance appraisal instruments in termination cases (i.e., *EEOC v. Zia Co. & Los Alamos Constructors*, 1976, 1978; *Harkless v. Sweeny School District*, 1975; *Powers v. Mancos School District RE-6 Montezuma County* 1976). This is not to imply that these issues are unimportant. From a theoretical perspective, continued research and advances in these areas help us to better understand what is undoubtedly a complex process. From a purely legal perspective, however, these issues seem to be accorded very little weight.

In a recent article, Banks and Roberson (1985) interpreted the performance appraisal process as a test development task. We see this line of reasoning as potentially damaging as it could eventually place the profession in a position it cannot satisfactorily defend. As discussed previously, it

is extremely difficult and perhaps impossible in many instances to support performance appraisals as if they were "traditional" tests. Moreover, the courts appear to be taking a more rational approach, which realizes that no matter how well developed or sophisticated a performance appraisal system is, it cannot prevent an individual rater from discriminating in his or her performance evaluations on one or more occasions. The key issue is obviously that of discrimination, not of refinement of format, performance appraisal development, or validation. Although in the context of promotions, instead of terminations, at least one court has explicitly stated that performance appraisals are not tests (*Mortensen v. Callaway*, 1982).

In the same year *Brito* was decided, the Supreme Court established the disparate treatment doctrine in *McDonnell Douglas v. Green* (1973). The disparate impact doctrine was established two years previously in *Griggs v. Duke Power Co.* (1971). Since these two decisions were fairly new, it is not surprising that *Brito* did not state under which theory the case was being decided. However, virtually all the cases reviewed subsequent to *Brito* also failed to state which doctrine (disparate impact or disparate treatment) was being applied. As most of these cases involved a single individual (or only a few), they appear to have been handled as if they were disparate treatment cases. While beyond the scope of the present review, it is important to note that courts differ as to whether a disparate impact analysis can ever be an appropriate method for testing the validity of subjective procedures such as the typical performance appraisal technique (Ashe & McRae, 1985). It could be argued that in a case where a large number of people have been terminated and adverse impact can be demonstrated, the subjective nature of a performance appraisal can provide the critical link between statistics and discrimination. Subjective appraisals do not have the same importance in disparate treatment cases as they may have in disparate impact cases.

The potential for supervisory biases to contaminate a performance appraisal is, of course, much greater with subjective approaches to evaluation. Consequently, courts do seem to consider whether these standards are evenly applied to all employees and whether organizations have a review process that mitigates the chances of individual supervisory bias. In addition, although not a necessary practice, the courts seem to react favorably to performance counseling aimed at correcting substandard work performance.

Since there is always the potential for error or deliberate discrimination in human evaluations, it is impossible to design an appraisal system that is completely "safe" from litigation (Burchett & DeMeuse, 1985). However, an employer can reduce the likelihood of being found guilty of discrimination and optimally defend against a discriminatory termination charge by following several guidelines. Our review of the court cases along with good professional practice suggests the following six points:

1. Conduct a job analysis to ascertain characteristics necessary for successful job performance.

2. Incorporate these characteristics into a rating instrument. Although the professional literature recommends rating instruments that are tied to specific job behaviors (i.e. BARS), the courts routinely accept less sophisticated approaches such as simple graphic rating scales and trait ratings. Regardless of method, written definitive standards should be provided to all raters.

3. Train supervisors to use the rating instrument properly. This involves instructions on how to apply performance appraisal standards when making judgments. The uniform application of standards is extremely important. In six of ten cases decided against the organization, the plaintiffs were able to show that subjective standards had been applied unevenly to minority and majority employees.

4. As demonstrated in several cases, formal appeal mechanisms and review of ratings by upper-level personnel is desirable.

5. The organization should document evaluations and reasons for the termination decision. This information may prove decisive in court. Credibility is enhanced with documented performance appraisal ratings and instances of poor performance.

6. Provide some form of performance counseling or corrective guidance to assist poor performers in improving their performance. As seen in several cases, the courts look favorably on this practice.

Although court decisions regarding terminations suggest that professional validity issues as applied to performance appraisal instruments are not of crucial importance, we certainly do not discourage efforts to properly validate and improve performance appraisal instruments. Organizations should be concerned about the integrity and fairness of their performance appraisal procedures. But, efforts designed to improve and validate performance appraisal instruments do not necessarily have the legal value that some commentators in the field believe.

REFERENCES

- Age Discrimination in Employment Act of 1967, 29 U.S.C. Sections 621-634 (1976 & Supp. III 1979).
- Ashe RL, McRae GS. (1985). Performance evaluations go to court in the 1980's. *Mercer Law Review*, 36, 887-905.
- Bailey v. MCI Telecommunications Corp., 29 FEP Cases 1457 (D.D.C. 1982).
- Banks CG, Roberson L. (1985). Performance appraisers as test developers. *Academy of Management Review*, 10, 128-142.
- Bass B, Barrett GV. (1981). *People, work and organizations: An introduction to industrial/organizational psychology*. Boston: Allyn & Bacon.
- Bennett v. Eggers, 26 FEP Cases 1408 (D.N.J. 1981).

- Bernardin HJ, Beatty RW, Jensen W. (1980). The new Uniform Guidelines on Employee Selection Procedures in the context of university personnel decisions. *PERSONNEL PSYCHOLOGY* 33, 301-316.
- Bernardin HJ, Buckley MR. (1981). Strategies in rater training. *Academy of Management Review*, 6, 205-212.
- Bledsoe v. Wilker Bros., 33 FEP Cases 127 (W.D. Tenn. 1980).
- Blunt v. Marion County School Board, 515 F.2d 951 (5th. Cir. 1975).
- Bradington v. International Business Machines Corp., 360 F.Supp. 845 (D. Md. 1973), *aff'd*. 492 F.2d 1240 (4th Cir. 1974).
- Brito v. Zia Co., 478 F.2d 1200 (10th. Cir. 1973).
- Brodgen AE, Taylor EK. (1950). The theory and classification of criterion bias. *Educational and Psychological Measurement*, 10, 154-186.
- Buchanan v. American Petroleum Institute, 26 FEP Cases 466 (D.D.C. 1981); *aff'd*., 713 F.2d 864 (D.C. Cir. 1983).
- Burchett SR, DeMeuse KP. (1985). Performance appraisal and the law. *Personnel*, 62, 29-37.
- Carter v. Dialysis Clinic, 30 FEP Cases 1734 (N.D. Ga. 1981).
- Cascio WF. (1982). *Applied psychology in personnel management*. Reston: Reston Publishing Co., Inc.
- Cascio WF, Bernardin HJ. (1981). Implications of performance appraisal litigation for personnel decisions. *PERSONNEL PSYCHOLOGY* 34, 211-226.
- Civil Rights Act of 1964, Title VII, 42 U.S.C. Section 2000e.
- Connecticut v. Teal, 457 U.S. 440 (1982).
- Conteras v. City of Los Angeles, 656 F.2d 1267 (9th. Cir. 1981), *cert. denied*, 455 U.S. 1021 (1982).
- Cornelius ET, Hakel MD, Sackett PR. (1979). A methodological approach to job classification for performance appraisal purposes. *PERSONNEL PSYCHOLOGY* 32, 283-297.
- E.E.O.C. v. Murphy Motor Freight, 448 F.Supp. 381 (D. Minn. 1980).
- E.E.O.C. v. Zia Co. and Los Alamos Constructors, 16 FEP Cases 1717 (D.N.M. 1976); *aff'd*., 582 F.2d 527 (10th Cir. 1978).
- Evans v. Meadow Steel Products, Inc., 35 FEP Cases 1195 (N.D. Ga. 1984).
- Faley RH, Kleiman LS, Lengnick-Hall ML. (1984). Age discrimination and personnel psychology: A review and synthesis of the legal literature with implications for future research. *PERSONNEL PSYCHOLOGY* 37, 327-350.
- Feild HS, Holley WH. (1982). The relationship of performance appraisal system characteristics to verdicts in selected employment discrimination cases. *Academy of Management Journal*, 25, 392-406.
- Foster v. MCI, 555 F.Supp. 330 (D. Colo. 1983), *aff'd*., 773 F.2d 116 (10th. Cir. 1985).
- Gidion v. Hospital for Joint Diseases, 20 FEP Cases 923 (S.D.N.Y. 1979).
- Grant v. C & P Telephone Co., 35 FEP Cases 1397 (D.D.C. 1984).
- Griggs v. Duke Power, 401 U.S. 424 (1971).
- Gupta v. IBM, 23 FEP Cases 857 (D. Md. 1980).
- Harkless v. Sweeny School District, 388 F.Supp. 738 (S.D. Tex 1975), *cert. denied*, 434 U.S. 969 (1977).
- Harris v. Group Health Association, Inc., 662 F.2d 869 (D.C. Cir. 1981).
- Hobson CJ, Mendel RM, & Gibson FW. (1981). Clarifying performance appraisal criteria. *Organizational Behavior and Human Performance*, 28, 164-188.
- Holley WH, Feild HS. (1975). Performance appraisal and the law. *Labor Law Journal*, 26, 423-430.
- Howard v. Miller Brewing Co., 31 FEP Cases 850 (N.D.N.Y. 1983).
- James v. Commonwealth Department of Public Welfare, 35 FEP Cases 1794 (W.D. Pa. 1983).
- Johnson v. Olin Corp., 484 F.Supp. 577 (S.D. Tex. 1980).
- Jones v. The Pitt County Board of Education, 528 F.2d 414 (4th. Cir. 1975).

- Kleiman LS, Durham RL. (1981). Performance appraisal, promotion, and the courts: A critical review. *PERSONNEL PSYCHOLOGY* 34, 103-121.
- Kresyman v. Bolger, 35 FEP Cases 665 (D. Kan. 1984).
- Landy FJ, Farr JL. (1980). Performance rating. *Psychological Bulletin*, it 87, 72-107.
- Lanphear v. Prokop, 703 F.2d 1311 (D.C. Cir. 1983).
- Latham GP, Fay CH, Saari LM. (1979). The development of behavioral observation scales for appraising the performance of foremen. *PERSONNEL PSYCHOLOGY* 32, 299-311.
- Latham GP, Wexley K. (1981). *Increasing productivity thru performance appraisal*. Reading: Addison-Wesley.
- Lilly v. Harris Teeter Supermarket, 720 F.2d 326 (4th Cir. 1983), *cert. denied*, 104 S. Ct. 2154 (1984).
- Marshall v. Kirkland, 602 F.2d 1282 (8th. Cir. 1979).
- Mason v. Pierce, 39 FEP Cases 21 (7th Cir. 1985).
- McDonnell Douglas v. Green, 411 U.S. 792 (1973).
- Mortensen v. Callaway, 672 F.2d 822 (10th Cir. 1982).
- Nath v. General Electric Co., 438 F. Supp. 213 (E.D. Pa. 1977) *aff'd.*, 594 F.2d 855 (3rd Cir. 1979).
- National Labor Relations Act, 29 U.S.C. Sections 151-168 (1935).
- NEA v. South Carolina, 434 U.S. 1206 (1978).
- Nord v. U.S. Steel Corp., 758 F.2d 1462 (11th. Cir. 1985).
- Planells v. Howard University, 32 FEP Cases 336 (D.D.C. 1983).
- Powell v. Syracuse University, 580 F.2d 1150 (2nd. Cir. 1978), *cert. denied*, 439 U.S. 984 (1978).
- Powers v. Mancos School District RE-6 Montezuma County, 539 F.2d 38 (10th Cir. 1976).
- Richards v. IBM, 504 F.Supp. 1369 (E.D. Mich. 1981).
- Schreiber NE. (1983). Wrongful termination of at-will employees. *Massachusetts Law Review*, 68, 22-35.
- Sias v. G.E. Information Services Co., 25 FEP Cases 1319 (D.D.C. 1981).
- Spiva v. Copperweld Steel, 22 FEP Cases 900 (N.D. Ohio 1980).
- Stieber J. (1984). Most U.S. workers still may be fired under the Employment-At-Will Doctrine. *Monthly Labor Review*, 107(5), 34-38.
- Stone v. Xerox, 685 F.2d 1387 (11t Cir. 1982).
- Taylor RL, Wilsted WD. (1974). Capturing judgment policies: A field study of performance appraisal. *Academy of Management Journal*, 17, 440-449.
- Texas Department of Community Affairs v. Burdine, 450 U.S. 248 (1981).
- Turgeon v. Howard University, 571 F.Supp. 679 (D.D.C. 1983).
- Uniform guidelines on employee selection procedures. (1978). *Federal Register*, 43, 38290-38315.
- Walston v. County School Board of Nansemond County, Virginia, 492 F.2d 919 (4th. Cir. 1974).
- Weahkee v. Perry, 587 F.2d. 1256 (D.C. Cir. 1978).

Copyright of Personnel Psychology is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.