Performance Appraisals: Developing a Sound Legal and Managerial System

Stanley P. Hodge

Largely because their development and application as functional management tools have been profoundly affected by legislative and judicial decisions, performance appraisal systems have recently become a primary focus of attention in human resource management. The current performance appraisal situation is reviewed in light of recent equal employment opportunity requirements and managerial developments. Guidelines and recommendations based on current criteria are provided in order to assess a library's existing performance appraisal system or as an aid in developing one that meets current standards. This paper also provides a tested method for developing an acceptable appraisal instrument for library classified staff.

Performance appraisal (PA) systems are often viewed by library managers as a necessary evil. There may be several reasons for this: the evaluation process may be considered as a chore that takes time away from more productive activities; employees may feel the process has little or no effect on quality of performance; and managers or staff may be dissatisfied with the appraisal instrument itself, either because it is inappropriate for the given situation or incorrectly administered.

An effective performance appraisal system involves more than just a rating form. It includes such factors as an evaluator’s judgments, job standards and criteria, organizational policy, legal requirements, and evaluator training. When properly developed and administered, a PA system can overcome many of the familiar criticisms and provide library management with a useful tool that may perform many functions in personnel decision making and improve employee effectiveness as well.

While there is little supporting empirical data to indicate the extent to which organizational functions are served by PA systems, there is consensus on seven general functions for which they are often used. These are:

1. To assist in personnel planning;
2. To provide a basis for employment decisions, i.e. promotion, termination, merit pay, demotion, etc.;
3. To guide job development;
4. To provide performance feedback to employees;
5. To elicit feedback from the employee;
6. To serve as a basis for modifying or changing behavior;
7. To determine the need for training and coaching.

Because PA systems are often used as a primary basis for decision making in the personnel area and serve to link the employee behavior to organizational rewards, it is important that they provide an accurate reflection of job performance. When they do not, an organization not only subjects itself to charges of failing to comply with equal employment opportunity legislation, but also jeopardizes the
progress of its employees and the achievement of its organizational goals. Recent surveys have indicated that many organization’s PA systems lag behind applicable federal guidelines. 

Any rating instrument that is used as a screening device for employee decisions is viewed by the Equal Employment Opportunity Commission (EEOC) and the courts as an ’’employee selection procedure’’ and thus is subject to Title VII of the 1964 Civil Rights Act or precedents set by a number of federal court cases. At the very least, a library should review its performance appraisal system to determine that its effect has not discriminated against those groups protected by Title VII of this act. If adverse impact is shown by a plaintiff, the employer must show that its PA system is job related. When job relatedness cannot be demonstrated, ’’the court may render the employer liable for back pay, court costs, specific management training programs for and/or promotion of more female and minority employees as part of the settlement of the case.’’

The following discussion briefly reviews the established legal requirements pertaining to PA systems and some recommended criteria to use as a basis for developing a system that will serve managerial goals and avoid legal liability. In addition, the findings will be applied in a methodology to design a performance appraisal form for library classified staff.

** LEGISLATION AFFECTING PERFORMANCE APPRAISAL **

Title VII is concerned with discrimination in all conditions of employment on the basis of race, color, religion, sex, or national origin. Employment decisions include the training, rewarding, reassigning, promoting, demoting, retraining, and dismissal of employees. When performance appraisals used for making any of the above decisions result in adverse impact, they clearly fall within the purview of Title VII and subsequent sets of government guidelines on employee selection. The EEOC was created and given the power to bring suit against certain types of employers found to be in violation of Title VII. Those employers include federal, state, and municipal agencies, educational institutions, and any organization with more than fifteen employees who work for more than twenty consecutive weeks.

In 1970, the *Guidelines on Employee Selection Procedures* broadened the scope of EEOC’s power to enforce compliance with Title VII. For instance, the definition of test was expanded to include all formal, scored, quantified, and standardized techniques assessing job suitability when these are used as a basis for any employment decision. Under these guidelines, personnel decisions that result in adverse impact are subject to challenge by the EEOC when a disproportionate number of the minority or protected group is screened out. EEOC guidelines also require that employment practices, i.e., performance evaluations, be validated if any of the components are found to have an adverse impact on these protected groups.

Because the 1970 guidelines defined tests to include any and all formally scored, quantified, or standardized techniques used for selection and appraisal purposes, many organizations abandoned formal systems in favor of informal, intuitive procedures. The 1978 guidelines then redefined test to also include unstandardized, informal, and unscored appraisal procedures and were more specific than the 1970 version with regard to adverse impact, indicating that adverse impact should be calculated according to the ’’Four Fifths Rule.’’ Other ap-

----

*Adverse impact occurs when a substantially different rate of selection in hiring, promotion, or other employment decision results in a disadvantaged position for members of a protected group.

†The Age Discrimination Act of 1967 and its 1978 amendment parallels Title VII and prohibits discrimination against workers between forty and seventy years of age.
approaches used by the courts in assessing adverse impact include: (1) internal comparisons made between percentages of minorities employed in high- and low-level positions; (2) labor-market comparisons of the percentage of employed minorities with the percentage found in the general population; (3) evidence that an employer intentionally or unintentionally restricted members of a protected group; or (4) evidence that an employer continued to seek applicants for a position when a qualified applicant was rejected.\textsuperscript{11,12}

**LANDMARK COURT CASES THAT AFFECT PERFORMANCE APPRAISAL SYSTEMS**

Four landmark cases\textsuperscript{13,14} relating directly or indirectly to PA systems have had an early significant impact on current standards and requirements. In these cases, it was determined that some type of discrimination resulted from the defendants’ biased or unstandardized use of a selection or appraisal system. Four fundamental legal implications based on these court decisions resulted.

1. In *Griggs v. Duke Power Company* (1971), it was ruled that employment criteria that adversely affect a protected group must be shown to be job related.
2. In the decision of *Rowe v. General Motors Corporation* (1972), subjective criteria were suspect and ruled to be considered as only one component of an overall process.
3. In *Brito v. Zia Company* (1973), performance appraisals were considered tests, and subject to validation.
4. In the decision of *Wade v. Mississippi Cooperative Extension Service* (1974), the court ruled that the defendant used an appraisal system based on personal traits that are subject to partiality and to personal taste, whim, or fancy of the evaluator and rejected the performance appraisal validation because it was not based on formal job analysis.

In more recent rulings\textsuperscript{15,16} the courts considered performance appraisal instruments as if they were “tests,” and applied the Uniform Guidelines when evaluating their validity. Personnel specialists within both the public and private sectors have examined these and the numerous other cases involving discrimination charges that resulted from performance appraisal. While there is no guarantee that any rating instrument or PA system design will prove successful in an employer’s defense, certain steps may be taken by library administrators to develop a sound legal and managerial PA system that has a favorable chance of being successfully defended.

**SUGGESTED CRITERIA TO USE IN DEVELOPING/ASSESSING AND APPLYING A PERFORMANCE APPRAISAL SYSTEM**

In order to understand more fully the implications of this review, the findings are presented in a way that the librarian-supervisor may apply them in a practical situation. The following list summarizes the legislative, judicial, and managerial criteria that would constitute a strong foundation upon which a PA system might be built. The list is based on a review of the recent literature on the topic and reports the advisory findings of academicians, lawyers, and personnel specialists regarding how PA systems can meet the aforementioned criteria when they are developed, assessed, or when an established system is applied in practice. By using these criteria, it is possible to develop a PA system that would not only be more acceptable to the library administrator as a management tool but also to the courts as well in discrimination suits involving personnel decisions resulting from the application of a PA system. Likewise, the criteria might be used by plaintiffs as a principal basis for developing their arguments in a suit against an employer.

**Criteria for Developing/Assessing a Performance Appraisal System**

1. The system is devised using job analysis and the enumeration of critical elements defined in terms of job descriptions and annual performance goals. Job analysis involves describing the tasks, duties, and responsibilities associated with a job.
2. Employees are involved in setting
criteria based on "critical job factors."
3. Performance standards or requirements for both critical elements and other important job aspects are set, either separately or within position descriptions.
4. There is an absence or minimum of evaluation of personal traits, e.g., those that may permit substantial subjectivity by the supervisor.
5. Precise, unambiguous language is used throughout the appraisal form.
6. The weight of each measure in relation to the overall assessment is fixed if the appraisal involves various measures of performance.
7. When validation studies are required due to adverse impact, they are preceded by formal job analysis.
8. Training programs for managers and supervisors on conducting effective performance appraisals are completed by all managers/supervisors.

Criteria for Application of the Performance Appraisal System
1. Performance expected of employees is communicated and goals and objectives of the ratee's job are made clear in terms of behavior and the results to be achieved.
2. The ratee is advised of the purpose(s) of the appraisal.
3. At least two levels of supervisors review an appraisal before an evaluation is presented to an employee, particularly when it results in an "unsatisfactory" rating.
4. Persons completing the appraisal base their ratings on a personal knowledge of the ratee's performance and contact with the ratee.
5. Problems that may be hampering job performance are discussed with the ratee.
6. An opportunity is provided for the evaluatee to voice opinions during the appraisal process.
7. Procedures exist for employees who disagree with any aspect of an evaluation to appeal to higher management or a review committee.

WHICH FORMAT TO USE?
A review of the literature indicates that there are numerous types of performance appraisal systems used in libraries. Among the most popular are the Graphic Rating Scale, Management by Objectives (MBO), Written Essay, Behavioral Observation Rating Scale (BORS), Ranking, Forced Choice, and Forced Distribution methods. The EEOC has not specified that any rating instrument is safe from litigation, and no single system is necessarily advocated. (It is not the instrument or process that is illegal but rather the consequences of the process.) Each has its advantages and drawbacks, and a library is advised to develop one that meets its own particular needs as a managerial tool.

An example of how the librarian-manager may develop a PA instrument that would meet the above-mentioned criteria is described below. In this case, the objective was to design a performance rating form for library classified staff engaged in technical services work. Some specific objectives of this form to assist library's management were:
1. To aid personnel decisions, i.e., merit pay, promotion;
2. To assess the need for job development and further training;
3. To provide performance feedback to the employee.

The instrument selected as an example for development was the Behavioral Observation Rating Scale (See appendix A). This consists of a number of related behavioral statements that are grouped into categories. Employees are observed and rated on a five-point Likert scale ranging from unsatisfactory to outstanding to describe how well an employee demonstrates those behaviors. The categories or behavioral statements may be weighted and the results quantified for a total average score, although this is not essential. Only five ratings are used for each behavior item because research shows that there is little gained by adding scale values beyond five.

Since in this case the appraisal form was developed to assess classified staff performing several different functions, some statements may not always be applicable, and the supervisor is given the option to indicate so by not rating on some items. (For instance, some staff may not perform supervisory functions.)
APPLICATION: DEVELOPMENT OF A BEHAVIORAL OBSERVATION RATING SCALE

Description of Organization for Whom the Appraisal Form Was Developed

The technical services functions at the library for which the appraisal form was developed consist of acquisitions, circulation, interlibrary services, processing, and resource development. Eighteen librarians and eighty-five classified staff are employed in these activities. Of the eighty-five classified staff, twenty are library assistants who have supervisory functions. Classified staffs' salaries range from about $8,000 to $16,000 per year. Their educational level ranges from a high-school diploma through a master's degree. Although a high turnover is characteristic of the lower-level positions in some divisions, several staff have seniority of fifteen to twenty years. Ninety-one percent are women.

Methodology and Results

The technique described below was not difficult to carry out; however, the cooperation of staff who provided and ranked the critical incidents in their jobs was instrumental in obtaining a valid and reliable list of characteristic job-related behaviors. A frequently used job analysis technique for developing a BORS is to develop a list of critical or important incidents of behavior. Supervisors and those whom they supervise are asked to provide observations about the critical requirements of the job. Generally, effective incidents are requested before ineffective incidents. This is done so the participant does not jump to the conclusion that the information is being sought to demote or terminate an employee.

Survey questionnaires (appendix B provides an example) were developed to elicit responses from the three categories of personnel (librarians, library assistants, and clerks/secretaries) about what they thought were the most important, critical job-related behaviors for their own jobs and for the co-workers within their divisions. Librarians were asked to list important job-related behaviors for library assistants and clerks/secretaries; library assistants for themselves and the clerks/secretaries; and clerks/secretaries for only their own category. All were also asked to list examples of unacceptable job-related behavior.

A representative sample of 25 percent of the personnel was desired for the survey. This would consist of 25 of the 103 total positions. Positions selected to survey were based largely on the distribution of job levels within each division.

Permission to conduct the survey was first obtained from the library's assistant director, the head of personnel, and each division head whose staff were to be surveyed. Division heads were asked to select experienced personnel from their division whom they felt had both a good grasp of responsibilities and the ability to verbalize critical behaviors. The survey questionnaire was tested for clarity of purpose with one division. No problems in comprehending what information was being sought were encountered by those initially surveyed. The procedure in administering the survey was to gather each separate division's staff together, distribute the survey, and briefly describe what they were being asked to do and why. They were told that they were a "select group" whom their division head felt would be able to provide significant insight into what was important in fulfilling their job responsibilities. They were told to indicate important job-related behaviors that they felt made the difference between doing a superior or poor job. They were asked to read over the form and to ask any questions.

Twenty-three of the twenty-five survey forms were returned. Over 250 job-related behaviors were supplied by the respondents. Many of these were very specific and duplicative. In addition, the undesirable behaviors were converted to be consistent with desirable ones, i.e., "Employee is consistently late for work" was converted to "Reports to work on schedule."

The reason for following the above procedure was to adhere to what researchers in this field believe satisfies the require-
ments of the Uniform Guidelines to allow employees to participate in identifying the critical elements of their job. Also, to promote their acceptance by employees, it is wise to proceed participatively when appraisal systems are developed or revised. The rating scale is thus developed from a systematic job analysis supplied by employees, for employees. This method helps to minimize the possibility that the behaviors described are too vague or are inappropriate to the job. It also results in an appraisal instrument that is content valid.

Since library staff usually perform a large and diversified number of tasks, the assessment of each would not be practical. Consequently, the 255 observations were matched and synthesized into thirty-six more-general statements. With the assistance of the library's head of personnel, those thirty-six behavioral observations were further refined, and divided into eight general performance areas. This procedure helped to reduce the appraisal instrument to a manageable size and to structure it into logical performance areas that could facilitate the rating and counseling process. The general performance areas were: supervision, job knowledge, work habits, responsibility, quality and accuracy of work, relations with supervisor, human relations/cooperation, and social behavior. In addition, when the individual responses from the original 255 were divided into thirty-six behavioral observations, they were tallied as to whether they were mentioned by a librarian, library assistant, or clerk-secretary to ensure that a generally proportional distribution among the staffing levels was obtained.

Reliability and Validity

Two additional steps were taken to assure reliability and content validity. First, the list of thirty-six behaviors was distributed to all twenty-three subjects who participated in the survey. They were asked to check what they thought were the two most important behaviors in each category and also to indicate one in each category that they thought was the least important. (Relations with supervisor was an exception since there are only two behaviors in this category.) Reliability means "dependability," "stability," or "consistency." The purpose of this second survey was to test how consistent respondents were in indicating the degree of importance placed on the various behaviors. Would, for instance, those behaviors that were most frequently suggested in the initial survey still be ranked as very important now that respondents would see behaviors they may have initially failed to mention?

The results of the second survey confirmed the reliability of critical behaviors listed by respondents in the first survey. For example, "trains staff patiently and thoroughly and informs them of updated procedures" was initiated by ten respondents on the initial survey. On the second survey, when twenty-three persons sampled were made aware of this behavior, it was listed as being one of the two most important in the supervision category by fourteen respondents, and as the least important by only one. Hence, the reliability or consistency of this behavior is demonstrated for the BORS instrument. Another behavior, however, "determines and assesses job priorities and delegates them to appropriate staff" received a low reliability rating. Table 1 illustrates the behaviors arranged by rank order of reliability within each category for the general performance area of "work habits." When determining the behaviors to be included on the BORS, those with the lowest reliability would be subject to deletion in the final edited version of the appraisal form. (Appendix A lists the twenty-eight considered most important of the original thirty-six.)

Second, content validity of the instrument was further tested by selecting a representative group of job descriptions for ten of the survey participants. The characteristic duties and responsibilities and personal qualifications listed on each job description were compared to the behaviors on the instrument and matched when possible. Matches between performance requirements and the behaviors were pos-
TABLE 1
WORK HABITS BEHAVIOR ARRANGED BY RANK
ORDER OF RELIABILITY

<table>
<thead>
<tr>
<th>Item No.</th>
<th>1st Survey No. Times Mentioned</th>
<th>2nd Survey Most Important</th>
<th>2nd Survey Least Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>10</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

Item No.
14. Organizes work schedule and uses time efficiently.
11. Follows through on assignments.
12. Gives prompt attention to priority responsibilities.
16. Able to perform in absence of close supervision.
15. Adequately documents work so steps are not duplicated or omitted.
13. Maintains work area in a well-organized manner.

Possible on twenty-seven of the thirty-six items. The omissions could have been the fault of the job descriptions more than the critical behaviors listed by the employees. It should be noted, however, that those items categorized as "social behavior" are not the type of thing traditionally indicated in job descriptions. The results further confirmed the content validity already inherent through the procedure applied in gathering the behavior statements. (These had been described as being critical elements of the job and critical for job performance.)

Additional Factors in a PA System

To assist the library's management in meeting the three objectives of the performance appraisal, a series of questions is developed that require a narrative response (see appendix A). It is also recommended that the ratee be provided with an opportunity to agree or disagree in writing regarding the appraisal statements and to comment if needed.

The training and instruction provided to the supervisors and raters is integral to the appraisal system. Research has indicated these will improve the reliability and accuracy of performance ratings. In addition, the importance of a supervisor's review and counseling session with the library staff member cannot be overstressed. This "discussion provides the opportunity for clarifying any differences in perceptions concerning the employee's performance which cause the person to feel that the rating on a particular statement may not accurately reflect actual performance. It is highly recommended that any newly developed PA system be tested and evaluated, on a small scale, separately or concurrently with one already in use by the employer.

CONCLUSION

A library's performance appraisal system that fails to incorporate current standards may have several negative consequences. It may not only trigger litigation when adverse impact results against protected groups, but it may also impede employee and managerial effectiveness as well. This paper has provided some checklists that library administrators may utilize in evaluating their present PA systems or to develop alternatives that adhere to current standards. A procedure was demonstrated that can yield a job-related performance appraisal instrument adhering to recent recommendations of professionals in the field of human resource management. Although the example is specifically related to classified staff in library technical services, the methodology may be generalized to a wide variety and level of tasks in other organizations.
REFERENCES


7. Ibid., p.12334.

8. The EEOC recommends that appraisal instruments be content valid; that is, they reflect important dimensions of job performance. Content validity is concerned with representativeness and relevance of the behaviors described within a rating instrument in terms of the critical elements of the job. Content validity is also concerned with whether the instrument adequately measures the behaviors considered critical for effective job performance. The courts also stress the importance of this issue. See Dena B. Schneier, "The Impact of EEO Legislation on Performance Appraisals," Personnel 55:24-34 (July/Aug. 1978).


10. The "Four Fifths Rule" is a rule of thumb and not a legal definition used by the EEOC for determining adverse impact. Hiring and promotion rates for minority groups which are less than four-fifths or 80 percent of the rate for the group with the highest hiring and promotion rate are regarded as evidence of adverse impact. See D. A. Cutchin, Guide to Public Administration (Itasca, Ill.: Peacock, 1981), p.43.


APPENDIX A: BEHAVIORAL OBSERVATION RATING SCALE DEVELOPED FOR LIBRARY TECHNICAL SERVICES STAFF

<table>
<thead>
<tr>
<th>CODE:</th>
<th>1 Outstanding</th>
<th>2 Meritorious</th>
<th>3 Satisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 Needs Improvement</td>
<td>5 Unsatisfactory</td>
<td>N/A Not Applicable</td>
</tr>
</tbody>
</table>

SUPERVISION

1. Trains staff patiently and thoroughly and informs them of updated procedures.
2. Accessible to staff for answering questions and solving problems.
3. Handles problems impartially and provides those supervised with constructive suggestions.
4. Monitors and controls workflow in assigned area and anticipates problems.
5. Monitors quantity and quality of staff performance; evaluates and treats assigned staff fairly.

**JOB KNOWLEDGE**

6. Understands assigned responsibilities and their relationship to end product.
7. Learns and applies procedures and policies and knows where to find them.
8. Periodically reviews procedures and suggests changes/improvements.
9. Accurately interprets information available in work tools.

**WORK HABITS**

10. Organizes work schedule and uses time efficiently.
11. Follows through on assignments.
12. Gives prompt attention to priority responsibilities.
13. Able to perform in absence of close supervision.

**RESPONSIBILITY**

15. Readily accepts suggestions and is receptive to new ideas and methods of accomplishing objectives.
16. Willing to accept added responsibilities when required.

**QUALITY AND ACCURACY OF WORK**

17. Thoroughly investigates and attempts to solve problems before referring them upward for resolution.
18. Checks own work for accuracy and completeness.

**RELATIONS WITH SUPERVISOR**

19. Requests assistance when appropriate.
20. Maintains and fosters communication channels with supervisor.

**HUMAN RELATIONS/COOPERATION**

21. Conveys helpful, cooperative, and friendly attitude when dealing with library patrons and staff.
22. Interacts well with co-workers to perform assigned responsibilities.
23. Conveys a positive attitude toward work and co-workers.
24. Resolves problems with co-workers and patrons in a mature manner.
25. Participates in resolving divisional problems and contributes positive suggestions.

**SOCIAL BEHAVIOR**

26. Reports for work on schedule and has good attendance record.
27. Avoids excessive socializing or disrupting others with noise.
28. Avoids abusing telephone or office equipment for personal use.

**OTHER CONSIDERATIONS**

29. What are this employee's major strengths?
30. What specific steps may the employee take to improve performance?
31. What job training or development programs are recommended to help further employee's progress?
32. Other comments by supervisor.

---

**APPENDIX B: INSTRUCTIONS FOR LIBRARY ASSISTANTS**

Think about the specific jobs you and your clerks/secretaries perform. What are the most important job-related behaviors required by you and your staff in your division? Some may be general behaviors that are important for all within the division. Others may apply more directly to those just within your unit. A representative sampling of desirable and important behavior relating to jobs
within your division is requested. These should consist of simple, brief statements. Examples might be:
1. Checks and monitors quality of work for clerks under supervision.
2. Attempts to resolve problems with library users or staff before turning them over to a supervisor.
3. Checks work for accuracy before submitting it for further processing.
Advice: Try to begin the statements with a verb, i.e. knows, prepares, trains, delegates, completes, etc. This may not always be possible, however.
I. IMPORTANT JOB-RELATED BEHAVIORS FOR LIBRARY ASSISTANTS (list minimum of five):

II. IMPORTANT JOB-RELATED BEHAVIORS FOR CLERKS/SECRETARIES (list minimum of five):