SUBJECTIVITY IN THE PERFORMANCE APPRAISAL SYSTEM
OF A DATA PROCESSING COMPANY

THESIS

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Sheryl Ann Bierstedt, B.A.
Denton, Texas
August, 1982

An attempt was made to determine the presence of subjectivity in the appraisal system of a data processing company. Ninety-one clerks were given individual performance ratings by their supervisors, including an overall rating and ratings on seven performance dimensions. A multiple regression performed on these data resulted in a set of empirical weights. Supervisors were also asked to rank the relative importance of each of the seven dimensions to the clerk job. The mean rankings were regarded as apparent weights. A comparison of the empirical and apparent weights led to the conclusion that supervisors were not rating their employees according to what they said was important for successful performance, thus introducing an element of subjectivity into the system.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBJECTIVITY IN THE PERFORMANCE APPRAISAL SYSTEM OF A DATA PROCESSING COMPANY</td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>16</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
</tr>
<tr>
<td>Instruments</td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>19</td>
</tr>
<tr>
<td>Discussion</td>
<td>22</td>
</tr>
<tr>
<td>Appendices</td>
<td>26</td>
</tr>
<tr>
<td>References</td>
<td>30</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>1.</td>
<td>Empirical Weights and Ranked Weight of Performance Criteria</td>
</tr>
<tr>
<td>2.</td>
<td>Mean Rankings and Ranked Weight of Performance Criteria</td>
</tr>
<tr>
<td>3.</td>
<td>A Comparison of the Ranked Weight of Empirical and Apparent Weights</td>
</tr>
</tbody>
</table>
SUBJECTIVITY IN THE PERFORMANCE APPRAISAL SYSTEM
OF A DATA PROCESSING COMPANY

There is probably not another subject area in the personnel relations field that has received more attention than that of performance appraisal. The process has often been referred to by other names such as performance review, performance evaluation, performance planning, or salary review. Yager (1981), however, made a distinction between the performance appraisal, the performance review, and performance planning. He considered the performance appraisal to be "an interpersonal process that occurs regularly and frequently between supervisor and worker" (p. 129). On the other hand, the performance review is "a periodic summarization and review of preceding performance appraisals" (p. 130). Yager considered performance planning to be "a third and higher level of communication that grows out of the performance review" (p. 130) and involves establishing work goals and plans for improved job performance. Yager felt strongly that these were three separate and distinct systems and were also separate from ratings of potential, promotability and pay.

While Yager's distinctions were useful, the present author has chosen to maintain a broader view of performance appraisal. For the remainder of this paper, the term
Performance appraisal will refer to "the systematic description of strengths and weaknesses within and between employees" (Cascio, 1978, p. 316).

Although the topic of performance appraisal has begun only in the last twenty-five years to receive a great deal of attention in personnel and management circles, Brett and Fredian (1981) pointed out that formal systems for performance evaluation are neither new nor unique to modern America Business. St. Ignatius of Loyola developed a system, which is remarkably similar to many current performance appraisal approaches, shortly after he founded the Society of Jesus in the 17th Century. Congress, in 1842, passed a law requiring the heads of the government's executive departments to make yearly reports on the effectiveness and productivity of clerks, using the check list system.

While it is not clear which corporation was the first to implement a formal system for appraising employees, General Motors Corp. had a policy for senior managers as early as 1918. By the end of World War II, the movement toward formalized performance appraisal systems took strong root in the American business world (pp. 60-61).

The goals of a performance appraisal system may vary. Hayden (1973) classified the purposes, or goals, of performance appraisals into two major categories--administrative...
and employee development. Appraisals conducted for administrative purposes could benefit the management of a company and provide a basis for decisions on promotion, transfer, merit increases, and reduction in force. Appraisals performed for the purposes of employee development could benefit employees by enlightening them on standards of performance expected of them by their supervisors and also provide insight on where they stand with their supervisors.

McGregor (1957) reported that performance appraisals were designed to meet three needs, one for the organization and two for the individual:

(a) They provide systematic judgements to back up salary increases, promotions, transfers, and sometimes demotions or terminations;

(b) They are a means of telling a subordinate how he is doing, and suggesting needed changes in his behavior, attitudes, skills, or job knowledge; they let him know "where he stands" with the boss;

(c) They also are being increasingly used as a basis for the coaching and counseling of the individual by the superior (p. 89).

Schnake's (1978) idea of the purpose of performance appraisals was a bit simpler than those previously mentioned. He stated the purpose was one "to reward an employee for his performance over the previous year and to motivate him to do better in the future" (p. 29).
Binis (1978) considered the performance appraisal to be an excellent tool for motivating employees. He stated:

A good performer needs to be told in a somewhat formal situation that the job he or she is doing is sincerely appreciated. And an employee who is not performing satisfactorily can be told exactly what the problem with his work is and in turn may appreciate the attention paid him and try to improve his performance (p. 12).

Luck (1955) felt that performance appraisal could allow management the opportunity to ascertain the "relative worth of each employee's service" (p. 57). Concurrently, management could also discover ways to measure characteristics that affect job performance and could collect valuable information on social factors that influence an employee's work. Performance appraisals could also provide management with criteria by which to reward employees on a more equitable basis. Finally, Binis (1978) claimed that appraisals ensured timely intervention by identifying the worker whose productivity and/or proficiency is deteriorating.

According to Sauser (1980), the performance appraisal has had five broad purposes—feedback, personnel actions, development, placement, and responsibility. Feedback referred to informing an employee of what job behaviors are expected of him and how well he is meeting those expectations, while personnel actions involved decision on promotion, salary, layoff, and dismissal. Development
referred to identifying particular training needs or special talents in employees. Placement involved attempting to "place employees in jobs that best fit their capabilities and limitations" (p. 12). Finally, Sauser's goal of responsibility involved keeping both employees and supervisors "on their toes" (p. 12) as well as increasing the supervisor's sense of responsibility to his employees and the employee's sense of accountability for actions on the job.

Locher and Teel (1977) stressed the role the performance appraisal system can have in protecting an organization from charges of discrimination in discharges, promotions, and/or salary increases. They felt accurate, objective records of employee performance could be a company's best defense against charges of this nature.

Cascio (1978) mentioned five purposes of performance appraisals. Among them were (a) providing a basis for personnel decisions, (b) establishing objectives for training programs, (c) providing feedback to employees, (d) serving as predictors for promotion decisions, and (e) serving as criteria in personnel research, i.e., test validation.

The list of performance appraisal goals espoused by different researchers could extend ad infinitum; however, as is evident, the list would eventually repeat itself and most goals would reappear numerous times. In summary, it is generally agreed upon that a performance appraisal system
can benefit both an organization and its employees in several ways.

There are as many techniques currently in use for evaluating performance as there are goals for the appraisal system itself. Most researchers have agreed that there are many alternatives available, but have disagreed on the relative merit of each method. Regardless of which method is better, some techniques are more popular than others. Eight of the most frequently used techniques are briefly discussed below:

**Graphic Rating Scales**

Supervisor ratings have probably been utilized more than any other technique in performance appraisals. For example, Landy and Trumbo (1976) found that of all the validation studies published in the *Journal of Applied Psychology* between 1955 and 1975, 72% of them relied on some form of rating scale. The graphic scale has been the most widely used rating method (Oberg, 1972). Typically, a graphic scale assessed a person on particular factors along a continuum (Cascio, 1978; Oberg, 1972). The supervisor recorded his judgement of the factor in question on the continuum, which ranged from a high to a low degree (Haynes, 1978). Points or intervals on the continuum could be defined numerically or could be anchored by some type of qualitative or adjectival statement. The number of intervals on the continuum could vary, although it has been suggested that it not exceed nine (Landy & Farr, 1980).
A variation on the typical graphic scale has been the Behaviorally Anchored Rating Scale (BARS). The basic principal of the graphic scale was retained but numerical or adjectival anchors on intervals were replaced by behavioral anchors. Basically, behavioral anchors were "statements of activity or critical incidents describing job-related behaviors along selected performance dimensions" (Michaelree, 1979, p. 48). The issue of whether BARS are superior to traditional graphic rating scales has not been settled, although BARS does appear to have the advantage of using rigorously defined performance dimensions (Cascio, 1978; Landy & Farr, 1980).

**Essay**

Also called Free Written Statements by Haynes (1978), essay appraisals required the rater to write a paragraph or answer open-ended questions on an employee's performance (Haynes, 1978; Locher & Teel, 1977; Oberg, 1972). Areas typically covered by appraisals of this nature were the employee's strengths, weaknesses, potentials, and training needs (Haynes, 1978; Locher & Teel, 1977; Oberg, 1972).

**Critical Incident Reports**

This method required the rater to record actual instances of significantly good or exceptionally poor behavior of an employee (Haynes, 1978; Locher & Teel, 1977; Oberg, 1972). The examples should be recorded as they occur, preferably on a daily basis.
Checklists

The rater was provided with a list of statements describing employee behavior and "checks only those statements which accurately describe the performance of the employee being rated" (Locher & Teel, 1977, p. 246). Often the statements were assigned scoring weights during the design of the form (Haynes, 1978). These weights were unknown to the rater and completed forms were evaluated by a staff specialist (Haynes, 1978).

Forced Choice

In this technique, the rater selected from a set of terms or statements those that were most or least descriptive of the employee being evaluated (Cascio, 1978; Oberg, 1972). The terms in each set were matched so that they appeared equally favorable or unfavorable to the rater (Haynes, 1978). An overall rating was obtained by scoring the descriptors according to a predetermined weighting system (Cascio, 1978; Haynes, 1978; Landy & Farr, 1980; Locher & Teel, 1977; Oberg, 1972). Typically, scoring was completed by someone other than the appraiser, in order to keep the appraiser ignorant of the weights (Haynes, 1978; Locher & Teel, 1977; Oberg, 1972).

Employee Comparison

Also termed ranking (Haynes, 1978; Oberg, 1972), this technique basically involved rank-ordering employees from best to worst according to their relative performance
(Haynes, 1978; Locher & Teel, 1977). Two commonly used methods for ranking were alternation ranking and paired-comparison ranking. Alternation ranking involved choosing the highest ranking employee and the lowest ranking employee, removing both names from the list, and repeating the process with the remaining pool of employees (Cascio, 1978; Haynes, 1978; Oberg, 1972). Paired-comparison ranking consisted of simply pairing every employee with every other employee and choosing the better of the pair (Cascio, 1978). The person chosen most frequently was ranked at the top of the list and so on, with the person chosen least frequently at the bottom of the list (Haynes, 1978).

Results Oriented

Haynes (1978) referred to this method of appraisal as a goals or objectives method. The more popular term for it has been Management by Objectives (MBO) (Locher & Teel, 1977; Michaelree, 1979; Oberg, 1972). In any case, the method included setting measurable performance goals in advance and comparing the actual performance attained with these previously set gaols (Haynes, 1978; Locher & Teel, 1977; Michaelree, 1979). The goals could have been set by the employer, employee, or both (Locher & Teel, 1977).

Objective Data

"Objective measures attempt to quantify individual results" (Haynes, 1978, p. 18). Typically, these measures included production data (quality and quantity of work) and
personnel data (absences, tardiness, overtime, accidents, etc.) (Cascio, 1978; Sauser, 1980). This technique was not usually used alone in an appraisal, as it did not consider any behavior factors (Haynes, 1978).

Often, several of these methods have been combined into one system. Typically, the method(s) chosen provided sufficient information to meet the objectives of the appraisal program. Primary consideration has also been given to the amount of time needed to complete the appraisal or the cost effectiveness of one technique over another. Ultimately, the decision as to which technique to use has been left to the organization.

When implementing an appraisal program, one factor organizations have been aware of was that numerous problems could be encountered in the system, regardless of the type(s) of method(s) in use. Some of these problems could be satisfactorily resolved; others were more or less inherent in any appraisal process.

One problem often dealt with was the negative feelings most supervisors and employees have had toward performance appraisals. While both parties may have fully realized the political and practical implications of the formal appraisal, this awareness has only served to heighten their anxiety over the situation (Cascio, 1978). Supervisors tended to resent their appraisal duties when there were forced to use and defend an evaluation form which they felt was improperly
constructed (Sauser, 1980). In addition, Hamelink and Hamelink (1980) found that supervisors were aware of the fact that their own evaluations were in part based on the conduct of performance appraisals on their subordinates. Perhaps McGregor (1957) described the plight of the supervisor best when he said:

Managers are uncomfortable when they are put in the position of "playing God." The respect we hold for the inherent value of the individual leaves us distressed when we must take responsibility for judging the personal worth of a fellow man (p. 90).

The performance appraisal has been important to most employees because it could mean the difference between a stagnant career and a promotion or raise (Binis, 1978). In addition, each employee had the right to a fair and honest review (Binis, 1978). Often, however, employees have lost respect for the appraisal process as a result of both careless attitudes by supervisors and bungled reviews (Binis, 1978; Ferarra, 1975). Ferarra (1975) stated:

When a performance review is bungled, employees begin to view it as some sort of management game rather than a real appraisal tool. When every employee gets a pay increase and a pat on the back once a year, the process is a waste of everyone's time (p. 3).

Employees have also felt that their true performance was not adequately represented by the results of the
appraisal due to an inadequate form (Sauser, 1980). Haynes (1978) reasoned the inability to foster any change in the appraisal system could only lead to defensiveness and antagonism.

Another problem that also existed was lack of communication between supervisors and employees. One facet of this problem was that all too often employees have not been informed of what was expected of them on the job or how they would be evaluated (Hayden, 1972). Schnake (1978) added that supervisors have not always given advance warning to employees that something was wrong with their performance. Later, when the employees received their reviews, they were left confused and angry. Binis (1978) stressed that such problems could have been avoided by explaining to the employees what was expected of them; by promptly spot-lighting good work and correcting poor work; by showing and telling employees where improvement was needed; and by justifying any actions taken, either positive or negative, with factual documentation of the employee's performance.

A problem inherent in almost any system utilizing supervisory ratings was that supervisors were unconsciously making various types of errors in their appraisals and, hence, biasing the results. The most common of these errors were leniency, central tendency, and halo. Leniency errors occurred when a rater was either consistently easy with his ratings (rates everyone high) or consistently difficult.
(rates everyone low). Central tendency errors were made by supervisors who were reluctant to give ratings at either extreme and thus rate everyone as average. Halo errors occurred when a supervisor tended to think of an employee in general as good or bad and then completed the entire appraisal to reflect this impression (Haynes, 1978; Sauser, 1980). It has been suggested that training raters to recognize these errors could result in a reduction of their occurrence (Cascio, 1978).

Organizations have encountered problems if they attempted to combine discussions of performance with discussions of salary. Although these topics are not independent of one another, Hayden (1973) suggested that the issue of salary could dominate the discussion "to the extent that neither the supervisor nor the employee is in the proper frame of mind to discuss plans for needed improvement" (p. 606). The two issues should be discussed in separate sessions to ensure honesty and accuracy of feedback (Yager, 1981).

Subjectivity, or lack of objectivity, was a common problem and has had a tremendous impact on the validity of an employee's performance appraisal. Sartain, North, Strange, and Chapman (1958) stated:

Objectivity in performance appraisal refers to evaluation which is unbiased, fair, impartial, dispassionate, and without prejudice. . . .Objectivity is the characteristic of an individual which leads him or her to interpret,
evaluate, and respond on the basis of what is actually discovered rather than as a result of feelings, emotions, and the like (p. 424).

Preston and Wichlep (1979) reasoned that performance appraisal naturally involved human judgement and subjectivity since it concerned people evaluating other people. They suggested that human judgement in the appraisal process was not necessarily bad because "evaluators are expected to exercise good judgement. They are not expected to exercise biased judgement" (p. 12). Calhoon (1949) rejected this line of thought and felt that "objectivity . . . is a snare and delusion since supervisors are unable to reject the general impression they have of people" (p. 516). Michaelree (1979) concurred and stated, "Man is not by nature an objective animal, and the supervisor by nature even less so" (p. 49). He, however, did not condemn subjectivity. In fact, Michaelree added that supervisor subjectivity was the key to a successful appraisal. He concluded that "any system that isolates objectivity in appraisal has truncated the nature of employee evaluation" (p. 51).

Other researchers, however, have agreed that subjectivity was a problem in performance appraisal systems that needed to be minimized (Cascio, 1978; Landy & Farr, 1980). Cascio (1978) listed several causes for the problem. Among them were individual differences in raters, individual differences in ratees, and interactions between rater and
ratee. Although there could be numerous differences between individual raters, Cascio felt the two most significant were position in the organization and "extent and nature of the rater's opportunity to observe ratee performance" (p. 336). Differences in ratees which Cascio felt might affect performance ratings were age, race, sex, and job tenure. Sauser (1980) also suggested that factors such as appearance and personality could affect a ratee's evaluation. In reference to interactions between rater and ratee, Cascio (1978) and Landy and Farr (1980) found that raters tended to give more favorable ratings to same-race ratees. Landy and Farr (1980) also found that "rater sex and ratee sex do not appear to interact in their effects on evaluative judgement" (p. 79), but admitted that additional research was needed in actual work settings.

The information reviewed here has made it apparent that any performance appraisal system could have problems of one sort or another. In order to meet the goals of the appraisal system and to facilitate the overall growth and effectiveness of the organization, these problems must be identified and dealt with. After an organization has become aware of the different problems that may occur in its appraisal system, the next step would be to identify those problems that actually exist.

The purpose of this study was to determine the extent to which the problem of subjectivity was present in the appraisal
process of a data processing company. The specific question addressed was whether supervisors in the organization are rating their employees as they say they are (i.e., are they actually rating their subordinates according to a reported weighting system), or are they unconsciously altering the weighting system and thus subjectively biasing their appraisals?

Method

Subjects

The subjects were 91 employees of a data processing company, classified as clerks under the company's personnel classification system. These clerks occupied 33 different job titles and were employed in various departments across the company. All clerks were classified as non-exempt, were paid on an hourly basis and were subject to the same schedule of salary and performance review.

Instruments

Two rating forms were used in this study, both developed by the researcher for the purpose of the study. Form A (Appendix A) was used to obtain actual performance ratings of the 91 individual clerks from their respective supervisors. Form B (Appendix B) was used to obtain rankings of the performance criteria utilized in Form A, upon which the clerks had been rated. The performance criteria were previously established by the company and were in use in the appraisal
process throughout most of the company at the time of this study. All copies of Form A and Form B had an attachment (Appendix C) which listed and defined the performance criteria, according to the company's personnel policies manual.

Procedure

The clerk population of the company was identified by use of the employee information system in service at the company. First-level supervisors who maintained at least one clerk under their direct supervision were identified through organizational charts from each of the divisions of the company. The researcher then contacted each supervisor \((N = 25)\) by telephone and explained that she was conducting a review of the company's performance appraisal system and would like to enlist his or her help. The researcher arranged subsequent individual meetings with each supervisor for the purpose of having Form A completed by that supervisor. At the meetings, supervisors were instructed to complete the form as if actually conducting an appraisal of each employee listed. They were first instructed to give each clerk an overall performance rating from 10 (high score) to 1 and then to rate each clerk on the seven defined performance dimensions, again on a 10 to 1 scale. Supervisors were asked to not view the overall performance rating as a summation or mathematical average of the ratings on the seven criteria. Supervisors were allowed to write "DNA" beside any criterion they felt was inappropriate, and were encouraged to use each numerical
rating one time only. In order to ensure confidentiality, no employee names were ever written on the forms. Employees were referred to by job title only. Supervisors were allowed to pencil in initials of the clerks they were evaluating at the time of evaluation, but were asked to erase these initials before returning the completed forms. It was requested that the completed forms be returned to the researcher within one week of the meeting date.

Approximately four weeks after the last completed Form A was returned, the researcher again contacted each supervisor by telephone and arranged a final meeting with each supervisor for the purpose of completing Form B. At these meetings, supervisors were asked to rank the seven performance criteria used in Form A on a scale from 7 (high score) to 1, according to how important they saw each characteristic as being to successful performance in the particular clerk jobs under their supervision. In this situation, the supervisors were not evaluating employees, but rather were evaluating the job of clerk and were specifying the relative importance of each of the criteria to that job. Again, supervisors were given the option of writing "DNA" beside any criterion they felt to be inappropriate to the job and were instructed to use each numerical ranking one time only. It was requested that completed forms be returned to the researcher within one week of the meeting date.
Results

The data analysis in this study consists of three steps. The first step involves determining the empirical weights of the seven performance criteria (i.e., independent variables) on the overall performance rating (i.e., dependent variable) from the information obtained on Form A. Criteria rated as "DNA" by supervisors are designated as zero ratings and are included in the data analysis. A multiple regression using stepwise selection, reveals that four of the independent variables—Initiative and Self Reliance, Quality of Work,

<table>
<thead>
<tr>
<th>Performance Criteria</th>
<th>Partial Correlations</th>
<th>Empirical Weights(^a)</th>
<th>Ranked Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiative and Self Reliance</td>
<td>.68</td>
<td>.59</td>
<td>7</td>
</tr>
<tr>
<td>Quality of Work</td>
<td>.36</td>
<td>.22</td>
<td>6</td>
</tr>
<tr>
<td>Persuasive Ability</td>
<td>.29</td>
<td>.14</td>
<td>5</td>
</tr>
<tr>
<td>Reliability</td>
<td>.23</td>
<td>.13</td>
<td>4</td>
</tr>
<tr>
<td>Technical Ability</td>
<td>.18</td>
<td>NA(^b)</td>
<td>3</td>
</tr>
<tr>
<td>Interpersonal Skills</td>
<td>-.06</td>
<td>NA(^b)</td>
<td>2</td>
</tr>
<tr>
<td>Quantity of Work</td>
<td>.007</td>
<td>NA(^b)</td>
<td>1</td>
</tr>
</tbody>
</table>

\(^a\)Refers to standardized Beta weights.

\(^b\)Not applicable: Partial r did not meet tolerance level.
Persuasive Ability, and Reliability—account for enough variance of overall performance to be entered into the regression equation. Three of the independent variables—Technical Ability, Interpersonal Skills, and Quantity of Work—do not reach the specified tolerance level ($p > .05$) to be entered into the equation and hence no weights are computed for these variables. In Table 1, the seven independent variables, their respective partial correlations, weights for the first four variables, and the ranked weight of these variables according to the partial correlation are shown.

The second step of the analysis involves computing the apparent weights (i.e., what supervisors say is most important to the clerk job) of the seven criteria from the data on Form B. Criteria ranked as "DNA" by supervisors are designated as zero rankings and are included in the data analysis. Mean rankings of the importance of each of the seven performance criteria to the clerk job are obtained across the 33 clerk job titles examined through Form B. In Table 2, the average rankings of the criteria are presented according to their respective ranked weight. Although the mean rankings of several of the criteria are quite close in magnitude, it is clear that supervisors feel that Quality of Work is by far the most important characteristic. Likewise, Persuasive Ability is quite definitely the least important characteristic to the clerk job.
Table 2
Mean Rankings and Ranked Weight of Performance Criteria

<table>
<thead>
<tr>
<th>Performance Criteria</th>
<th>Mean Rankings</th>
<th>Ranked Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Work</td>
<td>6.18</td>
<td>7</td>
</tr>
<tr>
<td>Quantity of Work</td>
<td>4.58</td>
<td>6</td>
</tr>
<tr>
<td>Technical Ability</td>
<td>4.36</td>
<td>5</td>
</tr>
<tr>
<td>Reliability</td>
<td>4.30</td>
<td>4</td>
</tr>
<tr>
<td>Initiative and Self Reliance</td>
<td>4.15</td>
<td>3</td>
</tr>
<tr>
<td>Interpersonal Skills</td>
<td>3.52</td>
<td>2</td>
</tr>
<tr>
<td>Persuasive Ability</td>
<td>1.91</td>
<td>1</td>
</tr>
</tbody>
</table>

The third step of the analysis is a comparison of the ranked weight of each empirical and apparent weight. Table 3 presents this comparison. Although there is good to moderate agreement on several of the criteria (e.g., Quality of Work, Reliability, Interpersonal Skills), it is noted that the criteria found empirically to be the highest weighted characteristics (i.e., Initiative and Self Reliance) is ranked third from the lowest according to what supervisors say is most important to the clerk job. Similarly, Quantity of Work receives the lowest empirical ranking, while supervisors rank this characteristic as the second in importance. A Spearman Rank-Order correlation computed on the seven pairs
Table 3
A Comparison of the Ranked Weight of Empirical and Apparent Weights

<table>
<thead>
<tr>
<th>Performance Criteria</th>
<th>Empirical Weights (Ranked Weight)</th>
<th>Apparent Weights (Ranked Weight)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiative and Self Reliance</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Quality of Work</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Persuasive Ability</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Reliability</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Technical Ability</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Interpersonal Skills</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Quantity of Work</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

of ranks reveals little relationship between them (\(\rho = .11, p > .05\)). This low correlation is partly an artifact of the small number of sets of ranks, but it is also obvious that these two sets of ranks are quite dissimilar.

Discussion

A comparison of the ranked weight of the empirical weights and the ranked weight of the apparent weights of job performance criteria indicates that, for the most part, supervisors are not rating their employees according to what they say is important to successful performance. In other words, the supervisors are unconsciously altering the weighting system they reported as appropriate for the clerk
job and are thus subjectively biasing their appraisals. It should be pointed out that the disagreement is most apparent where agreement is most important--at the extremes. Those qualities found empirically to be the most important (Initiative and Self Reliance) and the least important (Quantity of Work) are ranked moderately low and quite high, respectively. Likewise, Persuasive Ability is given the lowest ranking by supervisors but has a moderately high empirical ranking. These findings lend support to the claims of other researchers that subjectivity in a performance appraisal system is a very common problem and one that should be brought to the attention of management (Cascio, 1978; Landy & Farr, 1980; Sartain, North, Strange, & Chapman, 1958).

The major implication of these results for management is that if the company wishes to implement an appraisal system that utilizes weighted criteria, the best means for establishing the weights would be to compute empirical rankings such as those found in this study rather than to ask all supervisors what they think is important to the job in question. Such a procedure would serve to reduce subjectivity and maximize the reliability of the system. These empirical rankings can then provide a basis for a weighting system for the criteria used in the appraisal. The reasoning behind this is that the supervisor's empirical rankings are a reflection of the basis on which he
is going to rate his employees, regardless of what he will report as being most or least important to the job. Of course, there will probably still be an element of subjectivity in the appraisal process, but management will have the satisfaction of knowing it has minimized this influence to some extent. The combination of the empirical weighting system with adequate training of supervisors on how to conduct fair evaluations would improve the appraisal system to an even greater extent. Fundamental to the training program would be an explanation of each of the criteria being used along with concrete examples of behavior demonstrating various levels of performance within each dimension. This would perhaps put all supervisors on an even level and thus benefit employees by ensuring fairer, more consistent and equivalent appraisals.

Along these same lines, it would also be wise to inform the employees of the basis for their performance appraisals. Although this does not directly relate to the problem of subjectivity, since that is the supervisor's problem, it does have an impact on the relationship between management and subordinates. Such a step might help to open lines of communication between employees and their supervisors, thus alleviating some of the negative feelings associated with performance appraisals. It might have been beneficial to this project to have questioned the subordinates of the supervisors employed in this study on what they felt was
important to their jobs and on what basis they thought their supervisors were evaluating them. This information could then have been compared to the supervisors' sets of rankings and the degree of agreement or disagreement would have been an indication of the relationship and communication between management and employees. Again, this is only indirectly related to the problem of subjectivity in the appraisal process.

This study has demonstrated that there is an element of subjectivity in the appraisal system of a data processing company. Utilizing the supervisors of all employees classified as clerks in the company, it was found that supervisors indeed do not rate their subordinates according to what they say is important to successful performance on the clerk job. The empirical ratings found in this study should provide a basis for a weighting system of criteria used in performance appraisals, and should thus reduce the biasing influence of supervisor subjectivity.
Appendix A

FORM A

Please complete the following for each designated employee under your supervision.

I. First give each employee one overall performance rating from one to ten, with one being the lowest and ten the highest rating.

II. Rate each employee on each of the seven characteristics, defined on the attached page, again on a scale from one to ten. If you feel a particular characteristic does not apply to a job, write "DNA" in the blank beside the characteristic.

<table>
<thead>
<tr>
<th>A. JOB TITLE</th>
<th>OVERALL PERFORMANCE RATING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QUALITY OF WORK</td>
</tr>
<tr>
<td></td>
<td>QUANTITY OF WORK</td>
</tr>
<tr>
<td></td>
<td>INITIATIVE AND SELF RELIANCE</td>
</tr>
<tr>
<td></td>
<td>RELIABILITY</td>
</tr>
<tr>
<td></td>
<td>INTERPERSONAL SKILLS</td>
</tr>
<tr>
<td></td>
<td>PERSUASIVE ABILITY</td>
</tr>
<tr>
<td></td>
<td>TECHNICAL ABILITY</td>
</tr>
</tbody>
</table>

ADDITIONAL PAGES PROVIDED FOR SUPERVISORS WITH MORE THAN ONE EMPLOYEE UNDER THEIR SUPERVISION.
Appendix B

FORM B

Please complete the following for each designated position under your supervision.

Rank the characteristics, defined on attached page, according to what you feel is most important to successful performance on that job. Rank them from seven to one, with seven being the most important characteristic and one the least important. If you feel that one of the characteristics does not apply to a particular job, write "DNA" in the blank beside the characteristic.

Please return the completed form as soon as possible to the Personnel Department. Thank you for your time and cooperation.

A. JOB TITLE______________________________

_______ QUALITY OF WORK
_______ QUANTITY OF WORK
_______ INITIATIVE AND SELF RELIANCE
_______ RELIABILITY
_______ INTERPERSONAL SKILLS
_______ PERSUASIVE ABILITY
_______ TECHNICAL ABILITY

 ADDITIONAL PAGES PROVIDED FOR SUPERVISORS WITH MORE THAN ONE POSITION UNDER THEIR SUPERVISION.
Appendix C

Performance Evaluation Criteria

Quality of Work - Does work meet or exceed normal standards with respect to:

- Accuracy, thoroughness, timeliness.
- Understanding organization and procedures to contribute efficiently to accomplishing the group's goals.
- Maintain satisfactory quality during peak loads.
- Produce satisfactory results on routine work as well as special projects which may seem more interesting.

Quantity of Work - Does output meet or exceed normal standards with respect to:

- Speed, consistency, time utilization, meeting deadlines.
- Managing the volume of work assigned, including peak loads.
- Completing normal work in reasonable time periods.

Initiative and Self-Reliance - Consider ability to:

- Take action when appropriate on own initiative
- Carry out assignments in manager's absence.
- Innovate and improve ways of accomplishing jobs.
- Offer help to others during slow periods.
- Resolve questions and produce satisfactory results when instructions are unclear.

Reliability - Consider ability to:

- Utilize good judgement in the use of personal absence.
- Consistently adhere to reasonable hours of work.
- Observe approved policies and practices.

Interpersonal Skills

- Be courteous to clients and others contacted.
- Resolve conflict and/or problem situations with others.
- Obtain cooperation from others.
- Resolve competing priorities and still maintain a good working relationship with those involved.

Persuasive Ability - Consider ability to:

- Influence clients or coworkers to do things.
- Sell ideas, not force ideas.
Appendix C--continued

**Technical Ability** - Does employee:

- Possess appropriate level of expertise to accomplish job.
- Make attempt to keep informed of changes in his/her area of expertise.
References


