

Results are appealing as appraisal measures, because they appear to represent “the bottom line” of work performance. But for results to be good performance indicators, they need to be identifiable as a reflection of the appraised employee’s work, and rarely are results not affected by external factors such as the market, general economic conditions, the quality of supplies, and the work of others in the organization.

Another drawback to using measures of work results in a performance appraisal is that they do not provide guidance for improvement and development. Simply being informed that mastitis is raging in the herd, you lost three contracts, you spent twice as much on maintenance, or you struck out in a ball game does you little good and usually presents little surprise. Most employees whose work is clearly reflected in results already know what those results are. What they really need is some information about why they achieved the results they did and what to do differently in the future in order to improve upon them.

For all the reasons indicated above, and because performance **is** essentially behavior, most personnel management professionals favor appraising performance in terms of observable behaviors.

Better Instruments for More Useful Appraisals

Perhaps the single most common defect in performance appraisal systems is weakness of the measuring instrument. All too often the instrument, or rating form, focuses on areas other than work performance and therefore increases the opportunities for rater biases to operate.

The appraisal instrument is a tool. Like any tool, a well-designed form is not everything, but it certainly helps. Its very structure affects (1) the accuracy of performance measurement, (2) the perceived fairness and acceptability of the appraisal system to workers, (3) the usefulness of performance appraisal for both administration and development, and (4) the legal defensibility of decisions based on performance differences or “merit.”

The Legal and Regulatory Angle

Substantial consensus has developed among courts of law and personnel management experts that the criteria of adequacy for an appraisal instrument are the same as those for any measuring device: validity and reliability. The potential price of using invalid or unreliable appraisals for administrative decision making is illustrated in countless legal cases. The following two cases established important precedents:

In Albemarle Paper Company versus Moody (1975), the company had established employee selection criteria that it found related significantly to evaluated performance on the job. When the court found that Albemarle's performance appraisal system did not focus on well-defined job tasks determined through a careful job analysis, it threw out the selection test that had been validated against performance ratings. The court ruled that a low test score was therefore not a valid basis on which to deny Moody a position he had sought. It ordered the company to compensate Moody for legal fees as well as all of the income that he would have earned if selected for the position.

In Brito versus Zia Company (1973), the employer was found to have discriminated when it laid off a disproportionate number of protected group members based on their previous performance appraisal scores. After examining the appraisal instrument, the court held that the ratings did not really reflect performance on important job tasks, but rather the "best judgments and opinions of supervisors." Such opinions were deemed an unfair basis for selective discharge, and the company was ordered to make costly amends to employees against which it had illegally discriminated.



Attributes of appraisal instruments defensible against discrimination charges are defined by the Equal Employment Opportunity Commission's "Uniform Guidelines," state departments of Fair Employment and Housing, and the courts. They suggest that performance measures be based on critical elements of the job performed and that they be both valid and reliable indicators of work. The 1978 Civil Service Reform Act goes further to require in the Federal Service that (1) employees participate in defining critical elements of their jobs, (2) employees be evaluated solely on the extent to which they fulfill these requirements—not in comparison to one another, (3) formal appraisals be conducted at least once per year, and (4) rewards be tied directly to rated performance. The model established by the act is not a bad one to follow even outside the federal government.

Validity and Reliability

Validity and reliability are two fundamental criteria for measures of anything, including work performance. *Validity* is the extent to which something really reflects what it purports to be. In terms of work performance, a valid measure is one that assesses behavior in terms of job duties or task requirements. A well-written job description, based upon careful job analysis, is the best foundation for a performance appraisal instrument. The description should clearly state major job duties in behavioral terms. For effectiveness of repair work, a valid measure