

## Performance Assessment



One Monday morning, Roger, the farm manager, was confronted by two irate tractor operators who accused Francisco, the foreman, of unfair behavior. During the weekend, Francisco had employed a young tractor driver, with little seniority, to apply pesticides. The more senior employees were furious because the foreman had assured them there would be no tractor work available. When questioned by Roger, Francisco admitted to lying to the tractor drivers about the availability of work. He defended the decision, however, by explaining that the more senior employees were uncooperative. The relatively new employee, Francisco had argued, could outperform both of the more senior men. Roger asked Francisco to communicate these feelings to the two senior tractor drivers. Francisco apologized to the men for lying. As he explained the performance issue, the two tractor operators became increasingly sullen. One of the men, red-eyed, asked why none of this had ever been shared with them before. Francisco agreed this would never happen again and that he would let his subordinates know how they were performing. While these two tractor drivers never became super achievers, they did improve their performance considerably.

Jess Gómez, Ag Labor Consultant  
Bakersfield, California

After employee selection, performance appraisal is arguably the most important management tool a farm employer has at her disposal. The performance appraisal, when properly

carried out, can help to (1) fine tune and (2) reward employee performance.

When it comes to rewarding accomplishments, organizational psychologists have often warned that

when performance assessments are tied to pay raises, employees may become defensive and less open to change. Yet, how can pay raise decisions be made if it is not through some sort of performance assessment?

What seems to be called for, then, are two different types of appraisals. One, which I will call the *traditional* performance assessment, may be effectively used to judge employees for pay raise decisions, validation of the selection process, and evaluation of the effectiveness of other human resource management efforts. This is the focus of this chapter.

In Chapter 7, we introduce the *Negotiated Performance Appraisal* (NPA) as a performance management and coaching tool. To be useful, the NPA needs to take place with sufficient anticipation, long before the performance assessment process takes place. While for some jobs six months may be sufficient; for others, multiple years may be required.

## WHY PERFORMANCE ASSESSMENT?

---

Employee selection, training and just about any cultural or management practice—such as the introduction of a new pruning method or an incentive pay program—may be evaluated in part by obtaining worker performance data.

The evaluation may provide ideas for refining established practices or instituting new ones. For instance, assessment data may show that a farm supervisor has had a number of interpersonal conflicts with other managers and employees. Some options include (1) paying more attention to interpersonal skills when selecting new supervisors, (2) encouraging present supervisors to attend communication or conflict management classes at the local community college, or (3) providing the supervisor one-on-one counseling, mentoring or coaching.

Data from performance assessments can also help farmers (1) plan for long-term staffing and worker development, (2) give pay raises or other rewards, (3) set up a coaching session, or

(4) institute discipline or discharge procedures.

For selection statistical validation purposes (Chapter 3), it is easier to evaluate performance data when large numbers of workers are involved. Useful performance data may still be collected when workers are evaluated singly, but it may take years to obtain significant data trends.

### Feedback data

Feedback may be *qualitative* or *quantitative*. Qualitative comments are descriptive, such as telling the shop mechanic you appreciate the timeliness and quality of her repairs. In contrast, quantitative feedback is based on numerical figures, such as the percentage of plant grafts that have taken. Some researchers feel feedback is particularly useful when workers have an achievement objective (see Sidebar 6–1).

## ACHIEVING A WORTHWHILE TRADITIONAL PERFORMANCE ASSESSMENT

---

Here are some key steps you can take toward achieving effective performance assessment—ones that can be used to validate the selection process, evaluate the motivational factor of a pay for performance system, or make decisions about pay raises or promotions:

- (1) Select what performance data to collect.
- (2) Determine who conducts the assessment.
- (3) Decide on a rating philosophy.
- (4) Overcome rating deficiencies.
- (5) Create a rating instrument.
- (6) Deliver useful information to employees.

### Select what performance data to collect

One way to classify on-the-job worker behavior is by considering the three Ps—*productivity* (what was done), *personal traits* (how it was done, conduct) and *proficiency* (skill).



*Productivity* can be measured in terms of specific performance accomplishments. Examples include reducing calf mortality, increasing yield of the alfalfa crop, or diminishing bruises in the nectarine harvest.

*Personal traits* such as motivation, willingness to take criticism, cooperation, initiative, and dependability may be considered. Personal trait ratings are useful, even though they sometimes say more about how supervisors get along with an employee than how well the employee performs on the job. Farmers are unlikely to want to reward achievements—no matter how excellent they are—if a worker only performs

grudgingly and after repeated admonitions.

When personal traits are considered as part of a performance assessment, specific characteristics should be related to the job. Often, a personal trait issue can be translated into an achievement. Instead of talking about worker dependability (personal trait), for instance, one may want to address how well an employee reports on assignment completions (productivity).

*Proficiency*—skill, knowledge, and ability—plays an important role in worker performance. When assessments address worker proficiency factors (e.g., AI skills for a herd manager, accuracy

---

*While employees may typically compare themselves to others, there is little to be gained by having the organization promote such comparisons. In a healthy organization one employee's success need not mean another's failure.*

## SIDEBAR 6-1

## Performance Feedback

Performance improved substantially (11 to 27 percent) in a number of settings when workers were given specific goals to achieve and received performance feedback. Two examples from the logging industry show how goal setting can work, one with the harvesting of timber and the other with truck drivers. In one study logger productivity increased 18 percent and absenteeism decreased with the setting of specific goals. Logger crews who had set their own goals tended to meet them more often than when goals were set by supervisors.

In a second study management felt truck drivers were not loading their vehicles to capacity. Drivers—fearing a fine from the Highway Department, or even losing their jobs—seldom loaded their trucks more than 58 to 63 percent of capacity. After goals were set to load trucks to 94 percent of capacity, there

were some striking changes. Within the first month, truckers were on the average achieving 80 percent capacity. Within three months, they were frequently surpassing 90 percent. The company saved an excess of \$250,000 in a nine-month period.

In these studies, management provided a work environment where employees would not be reprimanded for failing to meet a goal. The truck drivers apparently tested management at one point by reducing their percentage of loading capacity. Only after being assured of management's support did drivers increase their efforts again. The researchers felt (1) goals had to be challenging but achievable; (2) the importance of worker participation in goal-setting varied; (3) employees had to be provided with needed resources; (4) competition may be permitted but not officially fostered by the organization; and (5) employees must be competent, as motivation without ability is of little value.<sup>1</sup>

*The critical incident technique involves noting instances where workers react particularly well or poorly. Such as when a milker noticed elevated milk tank temperatures even if not part of her job description.*



and high reliability in quality control ratings for packing shed personnel), they help assure worker interest in overcoming deficiencies that may be blocking future performance or growth. A farm personnel manager may be appraised in terms of understanding labor management principles, knowledge of applicable labor laws, skill in conducting interviews, or ability to counsel employees, for instance.

In evaluations, farmers need to strike the right balance between productivity and personal traits. Jobs vary in the importance that can be attached to such factors. An equipment operator who spends hours preparing land, furthermore, has less need for teamwork than two milkers who work side by side. Over-emphasis on personal traits may increase compliance at the expense of both creativity and performance. But stressing achievement over personal traits may lead to a philosophy where the end justifies the means—no matter how dysfunctional or unethical the behavior.

## Determine who conducts the assessment

Input into the assessment of worker performance may come from many sources including the employee, co-workers, supervisors, subordinates, or even persons outside the organization. Ratings from multiple sources usually yield more reliable performance assessments.

*Employee.* While employee input plays a vital role in the NPA (Chapter 7), they may also contribute to the assessment process. Employees have a vested interest in making positive comments about their own performance, however, and can usually benefit from outside evaluation when it comes to performance assessments.

*Co-workers.* At times co-workers have a better grasp for a colleague's performance than the supervisor, but co-worker evaluations have a tendency to be lenient or overly harsh. Sometimes co-workers hope management will read between the lines and praise irrelevant or insignificant factors. Peer review is usually anonymous and several peers are involved in the evaluation. This anonymity, while often needed, can also lend itself to abuses.

*Supervisor.* Performance assessment data obtained from the immediate supervisor is the most common rating source. Supervisors are often in the best position to give workers an honest evaluation. The danger in supervisory evaluations is the substantial amount of power and influence wielded, often by the hand of a single rater.

*Subordinate.* From time to time subordinates may be asked for input into the evaluation of their supervisor. When subordinates have an input into their supervisor's evaluation, supervisors have been known to improve their interpersonal relations and reduce management by intimidation. Issues of anonymity and adequate sampling of subordinates may be important in traditional assessments. As with co-workers, anonymity can also lend itself to abuses. Supervisors often have to make difficult decisions that will not please everyone.

## Outside the organization.

Evaluations by outside clientele may be useful in instances when there is much personal contact with outsiders or when the person being evaluated knows more about aspects of the job than the supervisor.

## Decide on a rating philosophy

performance assessment data can also be classified according to whether employees are compared against others or are rated against a standard.

### Comparison against others.

Normally, when comparing employees against each other, a few employees end up at the top and a few at the bottom. The majority end up somewhere in the middle in what is known as a *normal distribution curve* (also known as "grading by the curve," see Figure 6-1). Where the employee is ranked depends on how a person performs in comparison to others.

The principal advantage of the comparison method is preventing raters from placing all employees in one category (for example, all superior). Two disadvantages—especially when very few workers are involved—include assuming (1) employees fall in a normal distribution (there may be four excellent performers in a group of five, or none in a group of six), and (2) there are similar differences in performance between two adjacent employees, for instance, between those ranked 1 and 2 and those ranked 4 and 5.

*Rating against a standard* permits a supervisor to classify employee

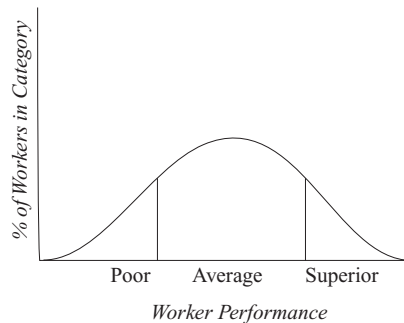


FIGURE 6-1

"Grading by the curve" (normal distribution bell curve).



*People's global self-esteem is affected by how they feel about specific areas of their life.*

performance independently from that of other employees. Both supervisor and subordinate have a reference point for accurately looking at an employee's long-term performance growth.

Ratings against a standard do not preclude comparisons. While employees may typically compare themselves to others, there is little to be gained by having the organization promote such comparisons. They are likely to create envy, vanity and dysfunctional

competition. In a healthy organization, one employee's success need not mean another's failure. If all can succeed, much the better.

Farmers who choose to use a standardized approach must next decide whether to judge all workers on an absolute standard or whether to consider an employee's time on the job. Those who prefer an absolute standard tend to give lower scores to employees, as they fear new workers who receive high marks will not feel the need for further improvement. In contrast, raters who feel a worker has done superior work considering his time in the position, may rate him as such. An evaluation six months or a year later yielding a superior mark would require a corresponding improvement on the part of the worker. The latter approach, which considers time on the job, seems more positive.

#### **Overcome rating deficiencies**

Supervisory evaluations often suffer from numerous rating deficiencies:<sup>2</sup> One particularly good or poor trait may contaminate other performance areas considered in the evaluation.

Once a worker is classified as a poor performer, it may take a long time for a supervisor to notice the worker has improved.

Supervisors tend to remember events more recent to the evaluation. Workers, realizing this, may strive to improve performance as time for assessments near.

Supervisors may tend to rate workers as average, especially when rating forms require a written justification for a high or low rating. Others may tend toward being either overly strict or lenient. Lenient raters may later appear to contradict themselves (e.g., when a worker is disciplined or does not get a raise):

“As with olives, where a small olive may be graded ‘large’ and the largest ‘super’ or ‘colossal,’ the worst rating many companies give their employees on appraisals is ‘good.’ Thus, the employer might be in the position of arguing that ‘good’ actually means ‘bad.’”<sup>3</sup>

## SIDEBAR 6-2

## Our Fragile Self-Esteem

We know very little about self-esteem despite all that has been written about it. Self-esteem seems to be somewhat situational, specific and fragile. Employees with very high or very low self-esteems (either end of the bell curve) may be less affected by the quality of their supervisors than those in the middle (the largest portion of the bell curve). Even so, few people can boast of self-esteems that are so robust that they cannot be deflated. A positive supervisor who looks for good in others is more likely to find it—as well as more receptive employees. Some supervisors, by their positive natures, seem to bring out the best in others.

Global self-esteem is affected by how people feel about specific areas of their lives, such as in their multiple roles as a spouse, a parent, a child, an employee, a supervisor, member of a team, and so on. Each of these general areas may be further broken down. A person, for instance, may feel she is generally a good supervisor. This same individual may recognize that she is better at some aspects of supervision than others. Furthermore, she may realize that her interactions with different members of her work team vary. If one aspect of our life is particularly important to us, our self-esteem in that area will have a large effect on our overall feelings of self-worth. Over time, what is important to an individual may change drastically.

Some suggest we need to focus on discovering and developing our talents. In many ways this is sound advice. Certainly, we do not have to be good at everything, and no matter how hard we try, this life is simply too short. To

really excel in an area can bring much positive satisfaction. It can also give us the confidence to venture into the unknown, or try something we are not good at.

A constant need to compare ourselves to others is a telling sign that something is amiss and that our self-esteem is weak. It is quite easy to confuse self-esteem with vanity and pride. In the quest for higher self-esteem and recognition people may ignore the truly important areas of their life. Individuals may even attempt to convince themselves that they can neglect critical areas of their life. Deep down, if not quite consciously, people know when their lives are out of balance, and this affects their self-esteem. Ironically, overly focusing on self can also destroy feelings of self-worth. It is often through service and building up of others that our own self-esteem is strengthened.

To excel at a job—one we have an interest in and talent for—requires we know how to appropriately focus our efforts. Trying to do more than we are able dilutes our efforts and little good comes out of it. An effective supervisor can help those under her analyze their efforts and find a positive balance. A healthy self-esteem will lower an individual's fear of identifying weaknesses and trying something new. Sometimes people resist learning or improving on skills that are essential to their jobs, but when they finally do make the effort, a sense of exhilaration comes into their lives that makes them wonder why they had not done so earlier. Such positive feelings may permeate all that a person is involved in, just as much as a feeling of stagnation does the opposite. If we stop growing, we stop living.

Raters may also be influenced by an employee's personal attributes such as national origin, level of education, union membership, philosophy, age, race, gender, or even attractiveness (Sidebar 6-3).

### Create a rating instrument

You can choose from several data collection and evaluation techniques, or rating scales. Whatever instrument is used, it should provide meaningful information to both employees and management.

Performance Area	0	1	2	3
Follows proper procedures to improve milk quality			☹	
Provides proper parlor environment for milking				☺
Recognizes and records cows in heat or sick			☹	
Keeps milk from fresh cows separate (cholostrum milk)				☺
Makes efficient use of time as cows are milked or washed				☺
Takes safety precautions with cows that kick		☹		
Cleans milking parlor for next milking				☺
3 = superior 2 = good	1 = below average 0 = not performed			

FIGURE 6-2

Numerical rating scale for milkers.

There are a number of ways of classifying performance assessment instruments. Data can be presented in terms of critical incidents, narratives, or predetermined anchors. A combination of approaches is often necessary to end up with a useful performance assessment. Assessment instruments require substantial rater training if results are to be meaningful.

**Critical incidents.** This technique involves noting instances where workers reacted particularly well or poorly. To be effective and accurate, critical incidents need to be jotted down as they take place and are still fresh in the supervisor's mind.

Examples of negative critical incidents include not observing elevated milk tank temperatures, or milking cows with antibiotics into the tank. Examples of noteworthy positive incidents are milkers who constantly provide accurate information on sick cows, or cows in heat; an employee who volunteers a money saving idea; or a worker who averted an upcoming disaster outside normal responsibility areas.

The strength of the process is in the concreteness of the examples provided. If care is not taken, though, the critical incident is susceptible to emphasizing negative worker behavior. When used alone, employees may have difficulty translating critical incident reports into improved day-to-day performance. Further, long periods of time may not yield any particularly good or poor behavior.

The critical incident approach can be used to come up with data and ideas to develop more complex rating scales.<sup>4</sup>

**Narratives.** As compared to the critical incident, narratives provide a

broader outlook on worker performance. Narratives work best when raters have the skills and take the time to provide a thorough, analytical report while maintaining a positive tone.

**Predetermined anchors.** Assessments where raters simply check or circle the most appropriate answer can potentially make for more standardized evaluations than either the narrative or critical incidents and are less time consuming for the supervisor (see Figure 6-2). Their ease in use may be deceiving, and raters may give the assessment less thought than it deserves. Anchor-based assessments include rating factors with a numerical scale (e.g., 0 to 3), or an adjective-descriptive scale (e.g., superior, good, below average).<sup>7</sup>

The most useful method is a combination approach that includes either a numerical or descriptive anchor, as well as critical incidents and a narrative performance description.

### Deliver useful information to employees

Assessments of worker performance work best when employees *know the evaluation criteria in advance*. In fact, in the NPA process employees help develop the goals and how these will be measured during performance assessments.

Such areas of evaluation can form the basis for an intelligent conversation about performance between supervisor and employee. In one farm operation a manager was able to not only discuss a foreman's performance within his present job, but also the types of skills that were needed if the foreman was interested in a potential promotion to assistant manager.

### SIDEBAR 6-3

#### Physical Attractiveness

Studies show attractive people are often judged to be more intelligent and have other positive qualities.<sup>5</sup> In one study, for instance, men gave attractive women higher scores on the quality of writing. Photographs of the supposed

authors were attached to the essays.<sup>6</sup> First impression attractiveness can have an even more serious impact on employee selection. This is particularly true where candidate impressions are formed solely on an interview and not moderated with data obtained from practical and written tests.



Despite the importance of formal assessments, an effective manager does not wait for these to communicate with employees. Sharing information about performance should be done frequently and in a positive manner. There should not be too many surprises for the employee when both discuss the evaluation.

When quality is an important factor of employee assessment, an excellent tool consists of providing feedback with performance cards which are handed to workers several times a day, based on quality evaluations (Chapter 11).

## JOB PERFORMANCE ASSESSMENT STATISTICS

---

There are a number of interesting statistics that may be employed for assessment of worker performance.

*Correlation coefficients* are useful for validation of employment tests such as those that measure vines pruned or pounds harvested within a specific time period. It is imperative that accurate records be kept, however. For instance, at one farm operation farm foremen would credit workers for the rest of the vineyard block in a given day, whether or not these were finished. Such a procedure made recordkeeping so much easier—but completely inaccurate. At others, when workers leave early, the exact departure time is not noted. Poor recordkeeping will result in unreliable productivity figures and lower validity correlations.

*Analysis of Variance.* My favorite statistical tool for evaluating worker motivation in piece-rate paid crews (Chapter 10) is the analysis of variance. For instance, we may examine twenty crew workers over a two week period in terms of daily performance. Performance data need to be converted into hourly averages if crews workers were employed for different hours each day. Through an ANOVA we may find out if there are statistical productivity differences *between workers* as well as *between days*. When crew workers are truly motivated, there will be a strong

statistical significance ( $p < 0.001$ , 0.01 or 0.05) between *workers*. There may or may not be statistical productivity differences between days, but these should be less significant than the differences between workers. Lack of statistical significance, low statistical significance, or higher significance between days than between workers all show—to varying degrees—lack of worker motivation.

*GageRR.* This quality control statistic (Chapter 11) permits us to continually test worker reliability and validity even after they have been selected for the job. We can thus evaluate pick vs. not pick (or pack vs. not pack) decisions. And when not packed or picked, we can evaluate how well individuals understand the reasons for not doing so (e.g., different types of fruit damage, fruit color, or fruit size). Pickers, packers, supervisors and quality control staff may all benefit from continual testing and feedback.

## SUMMARY

---

Key objectives of performance assessments include: (1) validating selection and other management or cultural practices; and (2) making decisions about pay or promotions.

Important steps to obtaining useful traditional assessments include determining the type of data to be collected as well as who will conduct the assessment, establishing a rating philosophy, overcoming typical rating deficiencies, creating a rating instrument, and engaging the employee in making decisions on future performance changes.

We also saw several types of statistics that can aid farm employers in the process of performance assessments.

## CHAPTER 6 REFERENCES

---

1. Latham, G. P., and Locke, E. A. (1983). Goal Setting—A Motivational Technique That Works. In *Contemporary Problems in Personnel* (3<sup>rd</sup> ed.) (pp. 313-324). Edited by Pearlman, K., Schmidt, F. L., and Hamner, W. C. New York: John Wiley & Sons.
2. For instance, see Carroll, S. J., and Schneier, C. E. (1982). *Performance Appraisal and Review Systems: The Identification, Measurement, and Development of Performance in Organizations* (p. 40). Glenview, Illinois: Scott, Foresman, and Company; Strauss, G., and Sayles, L. R. (1980). *Personnel: The Human Problems of Management* (4<sup>th</sup> ed.) (pp. 522-23). New Jersey: Prentice-Hall; and Kujawski, C. J., and Young, D. M. (1979). Appraisals of 'People' Resources. In *ASPA Handbook of Personnel and Industrial Relations* (pp. 4-193, 4-193). Washington D. C.: The Bureau of National Affairs.
3. Schlei, B. L., and Grossman, P. (1983). *Employment Discrimination Law* (p. 531). Bureau of National Affairs.
4. Carroll, S. J., and Schneier, C. E. (1982). *Performance Appraisal and Review Systems: The Identification, Measurement, and Development of Performance in Organizations* (p. 121). Illinois: Scott, Foresman, and Company.
5. For instance, see the work of Dion, K.; Berscheid, E., and Walster, E. (1986). What Is Beautiful Is Good, *Journal of Personality and Social Psychology*, (24: 285-290). ; and Bull, R. (1986). In Brown, Roger. *Social Psychology: The Second Edition* (p. 391-393). New York: The Free Press.
6. Landy, D., and Sigall, H. (1974). Beauty Is Talent: Task Evaluation as a Function of the Performer's Physical Attractiveness, *Journal of Personality and Social Psychology* (29: 299-304). In Brown, Roger (1986). *Social Psychology: The Second Edition* (p. 393-394). New York: The Free Press.
7. One popular performance appraisal approach of the "descriptive" type is the *behaviorally anchored rating scale* (or BARS). These scales are anchored with descriptive alternative behaviors. For every given category of behavior or performance, statements are ordered in an ascending or descending order of excellence. One challenge to BARS is the great number of descriptive category areas needed. Another difficulty is ordering observations so each statement of higher performance excellence subsumes the others.