14 BUILDING BETTER PERFORMANCE APPRAISAL SYSTEMS

Learning Objectives

- 14.1 Learn how organizations can reduce the negative consequences raters sometimes experience when they give accurate ratings
- 14.2 Understand why some changes to appraisal systems (e.g., changing rating scales) are often a poor use of the organization's time and effort
- 14.3 Learn some concrete strategies for building trust in performance appraisal systems
- 14.4 Understand how organizational dysfunctions can cause performance appraisal systems to fail

Performance appraisal is hard work, and some organizations seem inclined to simply abandon or severely curtail their appraisal systems (Buckingham & Goodall, 2015; Capelli & Tavis, 2016; Tziner & Roch, 2016). A number of high-profile organizations have made major changes to their appraisal systems (e.g., Buckingham & Goodall, 2015), and the business press has suggested that traditional performance appraisal is dying. The evidence suggests otherwise; performance appraisal is still with us, and there is no sign it is going to go away any time soon.

The great majority of organizations continue to rely on performance appraisal as an important tool for managing human resources (Lawler, Benson, & McDermott, 2012; Mercer, 2013). The question is not whether we can or even should get rid of performance appraisal, but rather whether performance appraisal can be improved. In Chapters 1–12 we outlined many of the challenges to evaluating performance in the workplace, and some of the problems that haunt performance appraisal (e.g., the strong motivation to give inflated appraisals) may be very difficult ones to solve. Nevertheless, we remain optimistic that performance appraisal can be substantially improved, and that the research that has been reviewed in this book points to some clear directions for improving performance appraisals.

It has long been argued performance appraisal research has not contributed very much to the quality of performance appraisals in organizations (Banks & Murphy, 1985; Ilgen, Barnes-Farrell, & McKellin, 1993; Murphy & DeNisi, 2008). There has been almost a century of research on performance appraisal, but this research has not led to notable improvements in the way performance is evaluated in organizations (DeNisi & Murphy, 2017). As a consequence, many consultants and practitioners seem to ignore this research and to advocate innovations that are not empirically supported, such as focusing on feedback or carrying out radical simplifications of appraisal systems. We think this is a mistake. Despite the somewhat mixed track record of using performance appraisal research to guide practice, we believe research on performance appraisal and performance management provides some eminently practical suggestions for improving performance appraisal in organizations. In this chapter, we lay out several ways of applying appraisal research to improve the practice of performance appraisal.

Improving Performance Appraisals

In Chapter 4, we laid out what we see as the four most serious challenges to evaluating performance in organizations. First, the task is difficult. Raters who want to accurately evaluate the performance of their subordinates must observe, recall, and integrate a lot of information, while at the same time attending to all of their other job duties. Second, the process and outcomes of performance appraisal are substantially influenced by a number of contextual factors, making it difficult to tell whether they represent assessments of performance or simple adjustments to the demands of the context in which they are carried out. Third, performance appraisals have many purposes in organizations, and these purposes can come into conflict. Finally, all of the participants in performance appraisal are motivated to distort the process in such a way that most employees receive high ratings. It is not easy to overcome all of these challenges, but we believe that performance appraisal research leads to some important and highly practical ideas for making progress on all four fronts. We group our suggestions for improving and reforming performance appraisal under four headings: (1) make appraisal easier and less risky for raters and ratees, (2) focus on the essential purpose of appraisals and sort out the conflicts among uses, (3) enhance the credibility and legitimacy of the system, and (4) do not waste time and effort on less essential features of your performance appraisal system.

Make Appraisal Easier and Less Risky

There are two reasons why performance appraisal research often has limited impact on the practice of performance appraisal in organizations. First, performance appraisal researchers have spent a lot of time studying topics that are of limited interest to practitioners (e.g., cognitive processes in rating; Banks & Murphy, 1985). It is also the case, however, that practitioners often fail to take advantage of findings that appear to have clear practical implications. Performance appraisal research *has* identified (and validated) a number of ways of making the task of evaluating performance easier and more reliable, and appraisals could be improved substantially if there was more consistent uptake of these suggestions by performance appraisal practitioners.

Two methods that could help make the task of evaluating performance in organizations easier have been extensively studied and validated. First, as we noted in Chapter 5, behavior diaries can be used to help reduce the demands on raters' memory when evaluating performance. There is evidence that the use of behavior diaries contributes to both the accuracy and the fairness of performance evaluations. These diaries require the investment of time and effort on the part of raters, and it would be naïve to expect that the contents of a behavior diary will be a completely accurate record of what raters have observed. The choice of what to record and what to omit from behavior diaries is influenced by a number of factors, and the contents of behavior diaries are probably biased by raters' overall impressions (i.e., raters who believe a particular employee is a generally good worker will probably attend to and record different entries in a behavior diary than would be recorded if the rater's overall impression was more negative; Maurer, Palmer, & Ashe, 1993; Murphy & Cleveland, 1995). Nevertheless, this method has some real potential, and we are struck by the lack of interest among practitioners in developing methods to make behavioral diaries simpler to use and more accurate.

Second, one of the real success stories in performance appraisal research has been the development of methods of rater training that are both successful in improving the consistency of performance ratings (Gorman & Rentsch, 2009; Woehr & Huffcutt, 1994) and viewed as reasonable by organizations and by raters (Sulsky & Kline, 2007; Uggerslev, Sulsky, & Day, 2003)—that is, frame of reference training. Some organizations provide this sort of training (Uggerslev et al., 2003), but we believe this method of training is not used in enough organizations, and that both the accuracy and fairness of performance appraisals could be improved if more performance appraisal practitioners took this method to heart.

The use of interventions such as behavior diaries and frame of reference training can make performance appraisal easier. The bigger challenge, in our view, is to make performance appraisal less risky for raters. In most performance appraisal systems, a rater who tries to give accurate performance feedback risks (1) harming his or her relationships with the employees who receive this feedback, (2) looking bad as a supervisor (i.e., the main job of a manager is to help his or her employees perform as well as possible, which means that low performance ratings are an admission of failure), and (3) getting no rewards from the organization for being accurate. As we have discussed in Chapters 9 and 12, the performance appraisal systems in most organizations create strong incentive to simply give everyone high ratings, and create strong disincentives to raters who attempt to give accurate performance feedback. Because raters and ratees are quite likely to have different views about the ratee's performance (see Chapters 5 and 9), it may not be possible to totally remove conflicts and differences in perspectives from performance rating, but we believe it is possible for organizations to take a number of steps that insulate or protect raters from some of the negative consequences of giving accurate performance feedback, by (1) making performance appraisal systems more transparent, (2) making them more fair, (3) making them forward-looking, and (4) reducing the visibility and impact of the direct supervisor's judgments.

Make the System More Transparent

Performance appraisal systems are more likely to be successful if the various participants have a clear and common understanding of the purpose of the system. Transparency is likely to pose especially daunting challenges in multinational corporations, where it is likely that both the actual purpose and the perceived purpose of performance appraisal will vary across units that are located in countries with different cultures and different political and legal systems (Maley, 2013). Nevertheless, increasing the transparency of appraisal systems is probably one of the first and easiest steps an organization can take to make the job of the rater easier and less risky.

Transparency has two components. First, the appraisal system must operate in a consistent and predictable fashion. Thus, one of the first tasks in making a performance appraisal system transparent is to remove or at least reduce the arbitrary and unpredictable parts of the system. This will reduce the flexibility and discretion of supervisors, managers, and executives, and transparency can potentially make performance appraisal systems bureaucratic and rule-bound, but this is arguably a price organizations must be willing to pay if they want to develop a system in which participants trust that it is their performance, and not the whim or self-interest of their superiors, that determines their evaluations. Second, communication is critical. If supervisors, managers, and employees do not *know* how the system works or how changes in employee performance are likely to affect things like rewards and sanctions, it hardly matters whether the system has achieved a high degree of consistency and predictability. Participants in the performance appraisal process act on their beliefs, not on the basis of objective characteristics of their environment, and the biggest challenge in achieving transparency is to get participants to *believe* that the system *truly* functions the way it is designed to function.

In the end, the communication problem organizations face is to convince the participants in the performance appraisal process that they will be treated consistently and reasonably. One way to accomplish this is to match actions to words—to actually and visibly treat employees in a consistent and reasonable fashion. For example, the U.S. Armed Forces represent one of the largest users of drug tests, and there was some initial concern that this type of testing could prove divisive and controversial. One of the decisions that led to a much higher level of acceptance of this type of testing was to make it uniform. *Every* active member of the U.S. Armed Forces, regardless of rank, is required to undergo a urinalysis drug test at least once a year (Smith, 2016). There are elaborate safeguards built into the system, but one key to its success is that it treats every member of the Armed Forces equally.

Transparency is a challenge to organizations for two reasons. First, transparency requires organizational decision makers to surrender power. In a truly transparent system, supervisors, managers, and executives cannot bend the rules or apply new criteria to shape the appraisal system to suit their purposes. Second, transparency requires careful and thoughtful design. A truly transparent system will probably make performance appraisal even more bureaucratic (it is not possible for everyone to follow consistent procedures and rules unless there *are* consistent procedures and rules), in part because a transparent system will require users to think through a wide range of contingencies. Suppose, for example, that you make a public commitment that anyone who receives ratings of "Exceeds Expectations" for all key performance areas will receive a raise. What are you going to do if all of the supervisors and managers in a department submit inflated ratings, giving the highest-possible ratings, even to employees who are visibly failing? A transparent system is going to require safeguards that protect both the users of the system and the organization.

Make Them Fair

We know a good deal about the factors that can make human resource systems seem to be either fair or unfair. Transparency can certainly contribute to perceptions of fairness, but transparency is often not enough to create the perception that an organization's performance appraisal system is fair.

There is a large body of research examining fairness and justice in organizations (Cohen-Charash & Spector, 2001; Colquitt, Conlon, Wesson, Porter, & Ng, 2001; Folger & Cropanzano, 1998; studies dealing specifically with performance appraisal include Erdogan, 2002; Flint, 1999; Holbrook, 2002; Korsgaard & Roberson, 1995). Most generally, this research suggests that three factors influence perceptions that a set of organizational practices is fair or just: (1) procedural justice, (2) distributive justice, and (3) interactional justice. First, the procedures used to evaluate performance and to use these evaluations to drive high-stakes decisions (e.g., raises, promotions) should be consistent, reasonable, and transparent (i.e., procedural justice). As we noted in the previous section, consistency and transparency are not easy to achieve; building an appraisal system that is viewed as procedurally just requires a good deal of thought and work. Second, the consequences of receiving positive or negative appraisals should be reasonable and consistent (i.e., distributive justice). That is, an employee who performs well should receive equitable, or at least consistent, rewards. Finally, the individuals who participate in performance appraisal should be treated with dignity and respect (i.e., interactional justice).

If you want your employees to believe that your performance appraisal system is fair, you must first give them some genuine reason to believe this, by making sure that the appraisal system *is* in fact fair, consistent, and reasonable. Beyond this, performance appraisal research has identified two additional courses of action that can contribute to the perception that your organization's performance appraisal system is fair: participation and voice.

Participation in performance appraisal is strongly related to satisfaction with and trust in appraisals (Cawley, Keeping, & Levy, 1998). Participation can mean many things, ranging from involving employees in the design and development of performance appraisal systems to two-way information sharing. In practice, many organizations involve employees in the development of their appraisal systems, and this involvement probably leads to increased acceptance of the system (Roberts, 2003). Giving employees voice—giving them opportunities to have input into their evaluations, voice their concerns, and have some expectation that their input will be taken into account—is a specific form of participation that appears to be particularly useful (Korsgaard & Roberson, 1995; Lizzio, Wilson, & MacKay, 2008; Whiting, Podsakoff, & Pierce, 2008).

Most generally, "voice" refers to behavior that changes the status quo with the intent of improving things rather than simply complaining or criticizing (LePine & Van Dyne, 1998). For example, employees who speak up about organizational issues or suggest ways of improving organizational processes are exercising voice (Van Dyne & LePine, 1998). In the context of performance appraisal, "voice" refers to giving the individuals who are affected by appraisals an opportunity to present information relevant to those appraisals. For example, many performance appraisal systems ask employees to rate their own performance or to comment on and sign off on appraisals, and these could both be thought of as forms of voice. Folger and Konovsky (1989) identified two distinct roles for voice in performance appraisal, feedback (i.e., providing information or input into the rating) and recourse (i.e., the ability to challenge or modify the evaluations you receive), and both appear to be useful. Voice appears to have both instrumental and non-instrumental effects (Korsgaard & Roberson, 1995). That is, voice can increase satisfaction with and trust in performance appraisals because the employee input often makes appraisals more accurate and more favorable to the employee (instrumental voice), but it can also contribute simply by virtue of the fact that an opportunity to provide input and to voice one's concerns makes the process seem fairer (non-instrumental voice).

One suggestion for increasing the effectiveness of voice is to build interventions to help make sure that the employee's voice is heard, taken seriously, and acknowledged. For example, many organizations ask employees to rate their own performance, but it is often unclear to the employee whether these ratings make any real difference. Suppose, for example, I rate myself as a "4" (out of 5) on Oral Communication, but my supervisor rates me as a "3." We think it is a very good idea for supervisors to (1) acknowledge this difference, and (2) provide some explanation for this difference. For example, a statement along the lines of "I see you rated yourself a '4' on Oral Communication and I rated you a '3.' Here are my reasons for assigning the '3' …" does several things. First, it acknowledges that employee's input. Second, it provides a rationale in cases where there is a difference in perspective. Of course, this sort of statement may open up a can of worms, in the sense that the employee might counter the reasons you offer for the rating of "3," but on the whole, we think this approach is worth the effort and risk. Voice is a good thing in performance appraisal, but if employees are asked to give input and come away believing that their input is ignored, the organization's efforts to provide opportunities for input could end up looking like a sham, and might lead to even *more* cynicism and distrust. If you are going to give employees voice, make sure that they know you are taking their input seriously.

Make Appraisals Forward-Looking

As we have noted in several preceding chapters, performance appraisal has many purposes, but two are usually cited as most important: (1) to support administrative decision such as raises and promotions, and (2) to provide information about strengths and weaknesses, which can be used to guide decisions about training. Unfortunately, many organizations devote meager resources to rewarding performance (Chapter 13 described the shortcomings of typical merit pay plans) and lack the resources to tailor training to the individual strengths and weaknesses of employees. One might argue that even in organizations that do not offer much in the way of merit pay or individually tailored training, there might still be some reasons for collecting information about job performance (e.g., for legal purposes, to use for validating tests), but it is hard to resist the conclusion that in traditional performance appraisal systems, too much attention is focused on evaluating past performance, and not enough is devoted to thinking about how to improve performance in the future.

A typical performance appraisal gives the individual and the organization fairly detailed information about how well each employee has performed over the last year. If the organization is unwilling or unable to do much in response to this information (e.g., because its merit pay budget is just too small to differentially reward good and poor performers), perhaps we should not concentrate so much on evaluating past performance. It might be better to devote more time and energy to the present and the future.

Many versions of performance management (e.g., Aguinis, 2013) involve giving employees frequent real-time performance feedback. This approach has two very attractive features. First, it focuses on the here and now—that is, giving feedback about what the employee is doing, and where necessary, helping him or her make changes that will improve subsequent performance. Second, because feedback is given in real time, it is likely to have the sort of immediacy and relevance that feedback about last year's performance may not possess.

Traditional performance deals largely with past performance. Performance management often focuses on the present. What about appraisal systems that are focused on the future? van Woerkom and de Bruijn (2016) suggest that instead of using retrospective appraisal systems, which focus on what you did (sometimes, focusing on what you did wrong) over the last year, it might be better to develop strength-based, forward-looking performance appraisal systems. They advocate developing appraisal systems that focus on identifying the unique abilities, skills, and other qualities each employee brings to the table and using those qualities to improve the organization. This sort of appraisal would shift the discussion from what was done well and poorly toward discussions of ways the employee might be able to contribute in the future toward accomplishing the mission of his or her work unit.

The feedforward interview represents an example of a future-oriented approach to performance appraisal (Bouskila-Yam & Kluger, 2011). In a feedforward interview (Kluger & Nir, 2010), employees are invited to tell about an experience at work during which they felt energized before knowing the results of their actions. This type of interview follows a three-step protocol, illustrated in Table 14.1.

The feedforward interview has two outstanding characteristics. First, it attempts to identify experiences at work that are intrinsically satisfying for employees. Second, it focuses on practical steps the individual and the organization can take to increase the likelihood that their experiences will be repeated. This method also has the interesting characteristic that it focuses on positive events.

Table 14.1 Feedforward Interview Protocol

Elicit a success story—Have the employee tell you about a positive work experience, particularly one in which the positive feelings are generated by the work itself, not the rewards.
Describe a success code—Identify the conditions and events that made this experience possible.
The feedforward question—Ask the employee to consider ways they can accomplish similar conditions and experiences in the future.

Neville and Roulin (2016) note that in most feedback situations, negative feedback can swamp positive feedback, an observation that is consistent with a wide range of research in cognition and decision making that deals with the way we process negative events (e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Kahneman & Tversky, 1979). They suggest that one strategy for reducing defensiveness in reaction to negative feedback is to maintain a substantial ratio (3 to 1 or more) of positive to negative feedback. The feedforward interview, with its singular focus on positive events and experiences, might represent one way of accomplishing this sort of ratio.

Assist and Protect Raters

A supervisor or manager who decides to be honest and accurate in his or her performance feedback faces a number of risks. If you give low ratings to subordinates who are hoping for and even expecting to receive higher ratings, it is likely that this feedback will hurt your future relationship with those employees, perhaps *decreasing* their motivation to work hard and perform well in the future. As we described in Chapter 12, supervisors and managers have many reasons to inflate ratings, and surprisingly few to provide honest and accurate appraisals. In Chapter 13, we noted that one way to get more accurate feedback to employees is to reduce the isolation and risk of the individual rater. If your supervisor is the only person who evaluates your performance, your disappointment over lower-than-expected performance ratings will be focused on him or her. This disappointment might be reduced if evaluations were seen as the product of the judgments of a group rather than a single individual.

In the past, we have expressed doubts about the value of 360-degree feedback (Cleveland, Lim, & Murphy, 2007), in part because a system that provides employees with feedback from supervisors, peers, subordinates, customers, and others is likely to sometimes give employees conflicting information. It is almost certain that peers or customers will rate a particular employee differently on at least *some* aspects of performance than supervisors rate that same employee. After all, different ratings sources are likely to observe different behaviors and may apply different standards when evaluating the behaviors they observe. Suppose, however, you modified the typical 360-degree protocol. One possibility, as described in Chapter 13, is that you might collect information from many sources, but feed it back in the form of a single aggregated report. We are attracted to this idea because it helps to solve multiple problems with performance appraisal, but it is important to understand the costs and benefits of this idea.

Suppose an organization solicited performance ratings and feedback statements from peers, supervisors, and subordinates, then fed them back to employees in a single report that did not differentiate among data sources. You could, for example, provide employees with a weighted average of all of the ratings you have collected rather than providing them with separate reports reflecting feedback from peer, subordinates, and others. You might even consolidate all of the comments about each performance area (e.g., Oral Communication) into a single report that does not treat comments from supervisors differently than comments from other sources. This sort of system would not be the most efficient, because it would require even more data collection than typical performance appraisal systems, but it could present several advantages. First, an employee who received disappointing feedback would know that this feedback is not simply the opinion of his or her supervisor, but rather represents the consensus of a large and diverse group. This might make disappointing feedback easier to accept, or at least harder to dismiss. Second, a system like this would increase participation and voice, since it would almost certainly include the opinions of coworkers, in effect involving each employee in the evaluation of each of the members of his or her own work group. Third, this sort of appraisal system would reduce the exposure of the individual supervisor, because disappointing ratings can no longer be blamed on the presumed biases or shortcomings of the supervisor.

The biggest impediment to adopting a system of this sort is probably not inefficiency, but rather the implications of this sort of "crowdsourcing" for the power, discretion, and influence of supervisors and managers. As we have noted throughout this book, performance appraisal is, in the end, an exercise of power. People who are at a higher level in a traditionally structured organization have the power and authority to evaluate the behavior and performance of their subordinates. We purposely use the term *evaluate* here, because traditional performance appraisal is not simply an act of reporting what subordinates have or have not done but also an act of assigning value to their performance. The appraisal system described here would substantially diminish the power and discretion of the individual supervisor, by shifting some of the responsibility for evaluation from the supervisor to the group of peers, subordinates, customers, and others who are asked to provide input. Suppose, for example, the supervisor is convinced that a particular worker deserves an overall rating of 2 (out of 5), but others in the organization think that a rating of 4 is more appropriate. The logic of the system described here is that the supervisor will just have to accept that his or her judgment is being overruled. Even if supervisory ratings are given more weight than ratings from other sources, this sort of system will inevitably say that in cases where the rater and others disagree, the rater will not fully get his or her own way.

We recognize that surrendering the power to unilaterally evaluate one's subordinates[1] can complicate the role of supervisors, but it could also liberate them. Rather than functioning mainly as a judge of their subordinates, taking the task of judgment and evaluation out of supervisors' hands could free them up to serve as coaches rather than judges. Instead of being seen as a problem by subordinates who receive disappointing evaluations, the direct supervisor might be seen as the solution—that is, as a person who can help employees who receive negative evaluations to improve.

Focus on the Most Critical Purpose of Appraisal

Performance appraisal systems in many organizations seem almost built to fail, because they are simultaneously used for multiple conflicting purposes (Cleveland, Murphy, & Williams, 1989; Murphy & Cleveland, 1995). For example, human resource researchers have understood for over 50 years that using the same performance appraisal system for both administrative purposes (particularly pay administration) and developmental purposes creates a set of conflicting demands and expectations that can make appraisal less useful for *either* purpose (Meyer, Kay, & French, 1965), but most organizations continue to use the same appraisal system for both purposes. As we have discussed in Chapters 8 and 12, we think there is compelling evidence that organizations shouldn't try to use the same performance appraisal system for both administrative and developmental purposes.

Spotlight 14.1 if at First You Don't Succeed, Stop Giving the Same Feedback

As we have noted in several preceding chapters, performance feedback does not always lead to improvements in performance. Performance appraisal researchers often treat the failure to see meaningful improvements in performance after giving feedback as a failure of the performance appraisal system, but this might not be the right diagnosis.

Feedback is most valuable when employees are new to their job and do not know whether or not they are performing their tasks right. At that point, performance feedback is often new and useful information. As employees gain experience, it is reasonable to expect that many of them know how to do their jobs, and also know whether they are performing adequately. For experienced employees, poor performance is not likely to be due to a lack of knowledge; performance problems among experienced employees are more likely to reflect a lack of motivation or willingness to perform.

Probably the least effective feedback strategy is to tell an employee the same thing year after year; if this approach does not work the first or second time, the odds that it will work the third of fourth time strike us as low. This suggests one practical step raters can take before giving performance feedback: to review feedback that has been given in the past and to consider the changes (if any) in performance following that feedback. Poor performers are already biased in the direction of rejecting much of the performance feedback they receive (see Chapter 9). If the feedback is simply a rerun of last year or of the last few years, it is probably that much easier to tune out this year's feedback. Knowing what you have told an employee in the past and how well or poorly that feedback worked might turn out to be very helpful in shaping this year's performance feedback.

We think it is time for organizations to choose which of these purposes—administration of rewards or employee development—is more important, and to live with that choice. If you want to use performance appraisal as a tool for deciding who gets a raise or a promotion, you should not use it to assess training needs. If you want to use performance appraisal as a system for employee development, you should not use it for salary administration. It might be feasible to develop two separate appraisal systems, one designed for each purpose, but we think this would be a mistake. These separate systems will almost certainly lead to conflicting assessments, because they focus on two very different aspects of performance. A system designed for developmental purposes must reliably measure individual difference in overall performance levels, and rating profiles that have lots of peaks and valleys (i.e., clear patterns of strengths and weaknesses) will tend to produce average ratings that cluster toward the middle, making it difficult to rank-order individuals in terms of their overall performance. Ratings with few inter-dimensional differences make it easier to sort individuals on the basis of their overall performance levels, but ratings of this sort will not provide much information about strengths and weaknesses.[2] A choice needs to be made.

In making a choice between these two important uses of performance appraisal, we think it is necessary to consider at least two separate questions: (1) how important are decisions about rewards versus about development in this organization, and (2) to what degree is the organization willing and able to take action in relation to these two uses of performance appraisal. The second question strikes us as more important, because many organizations claim that things like merit pay or rewarding good performance are very important, but in fact put too few resources toward rewarding performance to make the system work.

On the whole, questions about the capability and will to reward good performance are probably easier to address than questions about the capability and will to engage in employee development. An organization is not truly serious about merit pay if it (1) puts very little money at risk in a merit pay pool, or (2) routinely puts up with performance ratings that do not meaningfully distinguish good from poor performers. Thus, in theory, all it takes to make a genuine commitment to merit pay is to actually put a significant portion of each employee's pay at risk (as we noted in Chapter 13, the amount needed to make merit pay work is often a lot more than organizations or employees are willing to put truly at risk) and the development of an evaluation system that forces raters to distinguish among ratees. Training can be a bit trickier to evaluate, because the desire to provide training and developmental opportunities may not be enough. First, the structure of the organization, the nature of the labor market, and/or the legal system might constrain the organization's ability to offer a range of developmental experiences. For example, many large organizations offer expatriate assignments or assignments to different divisions or offices as part of their process for developing managers and executives. The effectiveness of these strategies is not always clear (Bolino, 2007), but the very availability of this method of training and development depends on factors such as the size and complexity of the organizations, and some organizations may not be in a position to offer this type of developmental experience regardless of their wish to develop employees. Second, a thorough assessment of training and development is likely to identify some weaknesses that are easy to address and others that are very hard to address through the types of training and development programs available to organizations. If a needs assessment reveals that the principal shortcomings of the workforce are in terms of highly stable attributes like general cognitive ability or broad personality traits, training and development is probably not the right strategy; improving recruitment and selection is more likely to help in these cases.

Suppose an organization determines that it is not truly willing or able to do anything meaningful to reward good performance or is not in a position to offer meaningful training and development. We suspect that many workplaces fall into one or both categories. As we described in Chapter 13, many organizations that claim to offer merit pay in fact do little more than go through the motions, giving very similar pay to the best and the worst performers. In a similar vein, there are several reasons to believe that many organizations that claim to be committed to employee development are in fact unable to do much to develop their employees. First, about half of all employed adults work in firms of organizations that are classified as large (500 or more employees), but more than one-third work in firms that are small (fewer of 100 employees) or very small (fewer than 20 employees; Caruso, 2015).

Opportunities for growth assignments may be quite limited in small firms, and even larger firms may find it challenging to offer this sort of developmental assignment to some of their employees. Second, organizational training budgets are often constrained, especially during economically challenging times. If you want to evaluate an organization's real commitment to training and development, it is more useful to look at the resources the organization actually devotes to training than to consult the HR handbook. An organization that says that it is committed to developing its employees but devotes scant resources to doing so is making a misleading claim.

Organizations that are unable or unwilling to do something meaningful with the information they collect in performance appraisal should seriously consider dropping performance appraisal altogether. The fact that an organization devotes time and effort to performance appraisal is likely to suggest to employees that it will do something with the information they collect, and we suspect that much of the dissatisfaction with performance appraisal is a result of the failure of appraisals to have meaningful consequences. Sometimes, the things that constrain organizations from acting on performance appraisals are temporary (e.g., short-term economic trends) and employees will probably put up with the situation if they know that you will act on performance appraisals once the current crisis or constraint is resolved. In too many organizations, performance appraisals appear to be little more than a waste of time, and the belief that your organization collects but ultimately ignores performance appraisals can undermine even the best appraisal system.

Perhaps the most common criticism of traditional performance appraisal systems is that they are a form of paper punching that does not lead to any meaningful outcomes. In theory, performance management systems should be less vulnerable to this criticism. The underlying philosophy of performance management is essentially active, with a focus on identifying opportunities to improve performance and supporting that improvement through feedback, training, and the marshalling of the organizational resources needed to perform well (Aguinis, 2013; Pulakos, 2009). In practice, many performance management systems seem to fall short of this ideal (Pulakos, Mueller-Hanson, Arad, & Moye, 2015; Pulakos & O'Leary, 2011), but nevertheless, a key distinction between performance appraisal and performance management is that performance management systems are designed to *do* something with performance assessments and feedback, whereas performance appraisal systems sometimes collect a lot of information but ignore that information when making decisions about how to reward or develop employees. In an organization that fails to link performance appraisals to meaningful outcomes, performance appraisal does little more than set up false hopes (e.g., that good performance will lead to valued rewards) and force employees to participate in a time-consuming and tiresome charade.

It is always useful for organizations to give serious thought to the question of *why* they are conducting performance appraisals. If they cannot come up with a convincing answer, this is a clear indication that the organization needs to carefully evaluate its priorities and to engage in a systematic redesign of its human resource strategy. While it might be tempting for organizations like this to simply give up on performance appraisal, we think a much better strategy is to give serious thought to *why* they are not making better use of a human resource tool that can serve important purposes in organizations, and to find ways to improve and effectively use performance appraisal as a tool for improving individual and organizational effectiveness.

Enhance the Credibility and Legitimacy of the System

You might have noticed that most of the suggestions for improving performance appraisals discussed so far have little to do with the topics that have dominated performance appraisal research for most of the last century, such as scale development, criteria for evaluating ratings, ratings versus rankings, or subjective versus objective measures of performance (DeNisi & Murphy, 2017). In fact, with the exception of a brief mention of behavior diaries and frame of reference training, the suggestions for improving performance appraisal discussed so far have very little to do with improving performance *measurement*.

We do not want to create the impression that good measurement is unimportant, but it is essential to understand that performance appraisal is not the same thing as performance measurement (see Chapter 13). Performance appraisal involves evaluative judgments (e.g., someone has to decide whether an employee's performance is good enough), social influence (e.g., performance appraisal is always to some extent political; Longenecker, Sims, & Gioia, 1987; Murphy & Cleveland, 1995), contextual influences, and the application of a range of human resource strategies, and good measurement does not guarantee good performance appraisal. Nevertheless, organizations can benefit from applying many of the strategies that have been put forth to increase the measurement qualities of performance appraisals. Despite our emphasis on topics other than the accuracy of performance measures, we would argue that many of the strategies developed over the years for improving the reliability and accuracy of performance ratings and other measures used in appraisal *are* important, but often for a very different reason than you might think.

Consider some of the things you might do if your goal was to develop reliable and valid measures of job performance. You would probably start with a careful analysis of the job, to ensure that the performance dimensions measured by your rating instruments were in fact important to that job. You would probably develop scales that make it as clear as possible what each performance dimension actually refers to and what the organization means by "very good performance," "average performance," "bad performance," and so forth on each dimension (a number of the behaviorally based scales discussed in Chapter 5 would help here). You would train the raters to make sure they were applying consistent standards in evaluating their subordinates (e.g., by implementing frame of reference training). You might collect data from multiple raters, to minimize the effects of idiosyncratic rater biases. All of these strike us as good ideas that might increase the precision of your performance assessments. More important, all of these steps are likely to contribute to the perception, on the part of several stakeholders, that the appraisal system is carefully developed, relevant, and administered in a thoughtful way. In other words, all of the interventions described in this paragraph are likely to contribute to perceptions of legitimacy.

Consider, for example, the standards the law relies on in evaluating performance appraisal systems that are challenged on the basis of discrimination against members of some protected group. Courts seem to focus on two general themes in evaluating performance appraisal systems: (1) is there evidence that the system is designed in such a way that it is likely to provide reliable and job-related assessments? and (2) are there mechanisms in place to provide employees a chance to have input and to challenge evaluations they regard as unfair? (Werner & Bolino, 1997). All of the steps described above for improving the measurement quality of performance assessments will help in establishing the system's legitimacy in the eyes of the law.

Similarly, visible efforts to increase the reliability and job-relatedness of performance appraisal systems, *regardless of their actual effects on measurement accuracy*, are likely to contribute to the perception that these systems are at least designed and intended to provide good measures. We are not sure whether following "best practices" will have a real effect on the quality of measurement provided by performance ratings, but we are convinced that visible efforts to do a careful and thorough job in evaluating performance will pay off in terms of perceptions of legitimacy. If raters and ratees believe that their appraisal system is poorly designed and hastily put together, they are unlikely to invest much time or effort in evaluating performance, and equally unlikely to take performance ratings and performance feedback seriously. Organizations that make concerted and visible efforts to ensure the job-relatedness, reliability, and accuracy of evaluations are sending a signal that they take appraisal seriously and that they are willing to invest resources to make it work well. The ultimate payoff from interventions designed to improve the measurement quality of performance ratings may be in terms of enhancing the credibility and legitimacy of the appraisal system.

Do Not Waste Time and Effort on the Less Essential Features of Your Appraisal System

Finally, in designing or retooling a performance appraisal system, it is important to devote your time and effort to the essential features of your system, such as the credibility and legitimacy of the system, the way the system protects the rater, its transparency, and so on. Unfortunately, organizations often devote time and effort to aspects of performance appraisal systems that are much less important, in particular, rating scales and technology. Neither of these is necessarily trivial, but in the grand scheme of things, time and effort you devote to these peripheral concerns is time and effort you probably will not devote to the more essential features of performance appraisal.

Our advice regarding rating scales is simple: Define the performance dimensions as clearly as you can and include some behavioral examples in your rating scales. Beyond that, it is doubtful that fiddling with the rating scale will have much of an impact. A rating scale is simply a tool, and developing the perfect rating scale will probably do little to improve your performance appraisal system beyond what you can accomplish with reasonably clear rating dimensions and a few behavioral examples.

Making large investments in technology designed to simplify the performance appraisal process strikes us as a risky bet. Many organizations are investing in databases, apps, computerized systems for organizing performance appraisal information, and the like.[3] There is nothing wrong with using technology to simplify a task like performance appraisal, but technology is often used in an attempt to solve the wrong set of problems. Many of the technological solutions to performance appraisal appear to be aimed at solving the same problems that rater training and rating scale development tried to solve in the 1950s through 1980s—that is, the rater's *ability* to accurately assess performance. We have come to believe that supervisors and managers usually have a reasonably good idea of who is performing well and who is performing poorly in their organization. The real shortcoming in many performance appraisal systems is that raters and not *willing* to provide accurate performance ratings and feedback (see Chapters 9 and 12), and technology rarely provides a solution to this problem.

Trust: The Essential Currency of Performance Appraisal

Trust is absolutely essential to the success or failure of performance appraisal systems (Farr & Jacobs, 2006). Trust that supervisors and managers will evaluate performance accurately and fairly, that they can give feedback (even negative feedback) without suffering negative consequences, and that rewards will be given when people perform their jobs well and withheld when they perform their jobs poorly are each important determinants of the success of performance appraisal. There is a good deal of evidence that the level of trust users show toward their appraisal system is related to both the psychometric characteristics of ratings (e.g., reliability) and to satisfaction with the system and motivation to devote time and effort to appraisal (Bernardin, Orban, & Carlyle, 1981; Bernardin & Villanova, 1986; Kay, Meyer, & French, 1965; Longenecker & Gioia, 1988). Organizational climates characterized by low levels of trust present a serious challenge to the effectiveness of appraisal systems (Lawler, 1971).

Better Performance Appraisal Requires Better Organizations

One of the key themes of contemporary performance appraisal research is that the context within which performance appraisal occurs is critical for understanding the success or failure of performance appraisal (DeNisi & Murphy, 2017; Murphy & Cleveland, 1995). This is one reason why "one-size-fits-all" recommendations (simplify, give more feedback) are likely to fail in the long run. The success or failure of performance appraisal and appraisal systems depends to a large part on what else is happening in the organization and even in society. For example, a well-developed performance appraisal system might fail if the organization does not have the resources to reward good performers. A system that is technically sophisticated but that does not fit the culture (e.g., a system that rewards individual accomplishment in an organization or in a culture that emphasizes cooperation and teamwork) will probably not succeed.

The suggestions laid out above for improving performance appraisal systems also say a good bit about the types of organizations in which performance appraisal is likely to succeed or fail. In particular, we believe a climate of trust and respect is absolutely essential if you want performance appraisal to succeed. You can design an appraisal system to be transparent, equitable, accurate, and impactful, but if the people who use this system do not *believe* that they are going to get fair and accurate evaluations, or that the evaluations they receive will result in rewards they care about, the system will fail. We often think about performance appraisal as a means for improving the way an organization functions, but perhaps we have it backwards. If you want to build better performance appraisal systems, it is probably best to start by building better organizations.

We believe that dysfunctional performance appraisal systems are more likely to be a symptom of a larger dysfunction in organizations than to be a problem that can be solved by adjusting the surface features of the appraisal system. In an organization that treats its employees with suspicion and contempt, or one in which performance appraisal is seen as an exercise in office politics (Longenecker et al., 1987), minor adjustments, such as changing the items on an appraisal scale or with the frequency of feedback, strike us as virtually pointless. In our experience, performance appraisal is rarely a big deal, and even more rarely a problem in well-run organizations that are characterized by a climate of mutual trust and respect. Instead of getting rid of performance appraisal, we should focus on getting rid of organizational climates in which mistreatment and abuse run rampant, and in which employees are treated as disposable cogs in a profit-making machine. Improve organizations, and you will almost certainly improve performance appraisal.

Spotlight 14.2 Performance Appraisal in Dysfunctional Organizations

There is little doubt that many organizations are deeply dysfunctional, and it is hard to image performance appraisal working well in these organizations. For example, a number of technology-focused or information-based firms in Silicon Valley have been criticized in recent years for their "bro culture," an organizational culture that combines the worst aspects of fraternity life with the arrogance of sudden wealth (Chrisler, Bacher, Bangali, Campagna, & McKeigue, 2012). A bro culture is characterized by immature, misogynistic, male-oriented behavior in which the only rule is "boys will be boys," and in which misbehavior is not only tolerated but often encouraged. There have been numerous reports of the difficulties this type of culture has caused companies such as Uber (Minter, 2017), but this culture is not limited to high-tech startups; similar cultures have been reported in major banks and investment firms (Lewis, 1989) and the military (Sorcher, 2013). Media organizations, notably Fox News, have been described as having a culture that tolerates the harassment of female employees by powerful males (Carterucci, 2017), and it is easy to see how such a culture could undermine even the best performance appraisal system.

The “bro culture” is not the only manifestation of dysfunction in organizations. There is a substantial research literature dealing with the “dark triad” of personality traits, a combination of narcissism, Machiavellianism, and psychopathy (Paulhus & Williams, 2002). Individuals who exhibit these traits tend to be manipulative, dishonest, unpredictable, and power-hungry, and this pattern of behavior can cause a number of problems in the workplace. There is evidence that this pattern of personality traits tends to draw people toward (and sometimes to succeed in) managerial and executive roles (Boddy, 2006, 2010; Jonason, Slomski, & Partyka, 2012), and even in organizations whose culture is generally tolerant and humane, the presence of a small number of dark triad managers or executives can create a climate of fear and mistrust.

Perhaps we don’t even need a dysfunctional climate or darkly motivated managers to create a lack of trust in organizations. In many organizations where the culture does not seem particularly vicious and where managers and executives seem to have fairly normal personalities, there are still good reasons why employees might distrust managers. In too many organizations, employees are treated like commodities and are shown little respect or consideration. At the same time that employees are being laid off and pensions are being decimated (Jones, 2017; Miller, 2017), executives seem to be treating themselves very well, often receiving hefty bonuses at the same time that employees are being cut. Interestingly, there is evidence that the highest paid CEOs are among the worst performers (Cooper, Gulen, & Rau, 2016), and it is possible that the decision to give outsized rewards to top executives is a sign of organizational dysfunction, not a sign of success. Regardless of how top executives are paid, organizational culture almost certainly matters in determining the success or failure of performance appraisal. An organization that treats employees like cogs in a machine should not be surprised if those employees respond with cynicism, detachment, and distrust.

Can organizations develop a culture of mutual respect, decency, and fairness? They probably can accomplish some of these things, but it is clear that building a culture of mutual trust and respect is a long and difficult process (Galford & Drapeau, 2003; Lewicki & Bunker, 1996; Sims, 2000). Unfortunately, it seems much easier to lose a culture of respect and trust than to build one. For example, Galford and Drapeau (2003) describe how even benevolent behaviors, such as tolerating an employee who is incompetent, can be corrosive to trust because these behaviors frequently lead to inconsistent standards and the appearance of favoritism. We suspect that it is easier to build trust in organizations that are relatively egalitarian, and harder to build trust when there is a great deal of separation between employees and management, but there have not been compelling empirical tests of this proposition.

Six and Sorge (2008) outline six methods that are believed to be useful in building a culture of mutual trust in organizations; these methods are outlined in Table 14.2 and are described below.

Promote a Relationship-Oriented Culture

In organizations, we often interact with others in terms of roles, such as supervisor–subordinate, coworker, and staff member. Roles cannot and should not be done away with, but one of the keys to building organizational trust is to encourage a culture in which we interact with others as individuals, not solely as people defined by their role (e.g., manager, team leader, production employee). Six and Sorge (2008) suggest that one step in building trust is to establish a norm that says relationships between people are important and that showing care and concern for others’ needs is valued.

Most organizations have handbooks that trumpet their support for a relationship-oriented culture, filled with phrases like “We Bring Out the Best in People,” but it takes more than verbiage to build a relationship-oriented culture.[4] First, in a culture of this sort, roles will be important for some purposes (e.g., determining who is responsible for what when performing a task), but they will not sort people into distinct classes that differ in their value or worth. Organizations that maintain a strong and rigid separation among levels (e.g., having different areas, amenities, and so on for executives versus employees) or in which roles strongly define the behaviors people are expected to engage in may find it difficult to put interpersonal relationships front and center.

Table 14.2 Strategies for Building a Trusting Organization

Promote a relationship-oriented culture
Facilitate unambiguous signaling
Socialize newcomers into the desired culture
Develop and manage the competencies required to sustain trust

One indicator of a relationship-oriented culture is a set of values and norms that discourages the aggressive pursuit of self-interest in favor of pursuing the interests of the community or group. However, we should not confuse relationship orientation with collective values. In highly collective cultures (e.g., Japan), one might strive for the benefit of the group without showing similar attention to or regard for individuals. In a truly relationship-oriented culture, the group is less than the sum of its parts, and it is critical to interact with individuals rather than with people who are defined in terms of their role, even when the role is group member.

Signal Your Intentions

One of the more important insights of Six and Sorge (2008) is that a culture of trust depends substantially on the ability of individuals to signal that they can and should be trusted. In practice, this involves providing the people you interact with frequent, clear, and credible signals that you intend to act in a way that enhances their well-being. The authors suggest two strategies that can help individuals appropriately signal, and accurately recognize when others are signaling, the intent to act in ways that advance well-being. First, signaling is a skill, and skills often require both instruction (training) and practice to master. Six and Sorge (2008) describe some of the methods companies they studied use to train individuals to sharpen their skills in signaling their own intentions and in recognizing others' signals. For example, one of the organizations they studied provided structured coaching for employees to help them communicate their intentions more clearly. They also suggested that organizations should encourage informal, face-to-face meetings, in part because these give people an opportunity to practice signaling skills in a less threatening environment than in more formal settings.

This focus on signaling helps to bring home an essential point about the development of trust in organizations. It is not enough for individuals to *have* benign intentions, they must also communicate their intentions to others. Trust involves taking a risk—in particular, a risk that your effort to look out for the well-being of the person you are dealing with will not be reciprocated. Effective communication of intentions in a relationship can mitigate this risk.

Socialize New Members

Six and Sorge (2008) suggest that organizations need to devote time and resources to socializing new employees, with a particular emphasis on helping newcomers understand the norms and policies of an organization that support a relationship-oriented culture and helping them understanding the importance of developing the sort of signaling skills that make this culture possible. These authors emphasize the importance of being detailed and explicit in describing not only the values, but also the behaviors the organization prefers and expects. This type of socialization is especially valuable because statements of corporate values (especially mission statements or other similar documents) can be bland and abstract. Many organizations say things like "people are our most important resource," but unless you know what (if anything) this actually means in terms of day-to-day behavior, this value statement may not provide a sufficiently concrete indication of what you should actually do to put this value into practice.

In their paper, Six and Sorge (2008) describe two different organizations, one of which had a long and detailed socialization process, where new organization members were expected to learn and understand the norms and values of the organization before they were "released" into the organization. The other organization followed a culture of "throwing new members into the deep end," putting new members of the organization directly into their new job with little instruction on the precise demands of the job, much less the norms and values of the organization. Their conclusion was that the failure to socialize new members of the organization left many opportunities for miscommunication and misunderstanding and for behaviors that (unknown to the new member) were not received favorably by their coworkers. At least in the case of these two organizations, difference in socialization practice seemed to be manifest in differences in the level of trust shown within the organization.

Monitor and Develop Competency

Six and Sorge (2008) note that "trustworthiness consists of two dimensions, ability and intentions" (p. 877). Part of trust is a belief that the people you interact with intend to behave in ways that advance well-being, but ultimately, exhibiting these intentions in a work setting involves behaving in ways that allow and assist others in carrying out their job tasks and functions. From this perspective, competence is an important component of trust. That is, if you have the knowledge and skills required to do your job and others have the knowledge and skills required to do theirs, it is considerably easier to develop trust based on mutually supporting behavior than if both you and the person you are interacting with lack the competencies to perform well.

This emphasis on competence is one of the key differences in developing and maintaining trust in a work organization as opposed, for example, to an informal social group. In a relationship-oriented work organization, you are expected to behave genuinely and benevolently toward other individuals, all of whom are trying to perform tasks, complete projects, manage processes, and the like. In this sense, competence is an integral part of this benevolence and respect. Individuals who are not competent to perform their jobs well are not likely to be in a position to actively aid others in carrying out the core missions of the organization.

The Ethical Practice of Performance Appraisal

A long series of major frauds, crimes, and dishonest practices by organizations, ranging from Enron and WorldCom to Volkswagen, have reinforced the importance of ethical behavior in organizations. While much of the research literature in this area focuses on identifying and preventing major ethical lapses (e.g., white-collar crime, financial manipulation), ethical considerations are important in evaluating a wide range of organizational policies, including the human resource practices of organizations (Buckley et al., 2001). For job applicants and employees, HR represents one of their primary points of contact within larger organizations, and virtually every HR policy an organization might adopt has potential ethical implications. For example, in 2015, California adopted a Fair Pay Act aimed at eliminating pay gaps between men and women performing comparable work or work of equal value. This act is considerably more stringent that the Federal Equal Pay Act of 1963, which mandates equal pay for men and women who perform jobs that are essentially identical, but that does not require equal pay for similar work. The California law is arguably an outgrowth of an ethical judgment that the failure of many organizations to pay men and women on a comparable basis when the work they perform is essentially similar (but not necessarily identical) was fundamentally unfair.

Ethical judgments are required when implementing many human resource policies. For example, it is well known that cognitive ability tests are valid predictors of performance in many jobs, but that these tests are also most likely to systematically screen out members of many minority groups from employment opportunities (Murphy, 2002). This conflict between efficiency and equity is not one that is easily resolved; an ethical judgment is required to resolve the conflict between wanting to use the most efficient and valid tests versus providing fair opportunities for members of minority groups to compete for jobs.

Performance appraisal involves a number of ethical judgments (Buckley et al., 2001). For example, should employees have a right to provide meaningful input in the evaluation of their performance? Note that this is a different question than asking whether giving employees voice in performance appraisal is beneficial (to them or to the organization). It involves considering whether a fundamental respect for human dignity implies that failure to give employees voice is wrong, regardless of the effects of voice for them or for the organization.

Raters face a particularly challenging set of ethical judgments when they decide whether to give performance ratings that benefit the ratee, the rater, or the organization. (See Chapters 7, 8, and 13 for discussions of the way the interests of raters, ratees, and organizations can clash in performance appraisal.) For example, a rater who gives inflated ratings might benefit the ratee, by increasing the likelihood that the ratee receives a raise or a promotion, while at the same time giving the organization misleading information about the ratee's performance. It is useful to think through the ethical issues that might confront raters, ratees, and organizations as a result of their conflicting needs and interests when conducting performance appraisals.

Buckley et al. (2001) note that performance appraisal is one of the processes that frequently tests the norms and values of the participants in the appraisal process. Raters, who probably have a very good idea of who is performing well and who is performing poorly, have a number of reasons to give their subordinates high ratings rather than ratings that reflect their real performance levels (Murphy & Cleveland, 1991, 1995). Employees are strongly tempted to engage in impression management behaviors that might help them get a favorable rating, even if the impressions they strive to create are not a realistic reflection of their actual effectiveness. Human resource managers often find themselves claiming that performance appraisal is important, when they know full well that ratings are often collected then ignored.

In theory, people ought to be held accountable for their behavior (Tetlock, 1985), but performance appraisal represents a situation in which the deck is stacked against honest and accurate evaluations, or against honest ways of talking about human resource policies. We think this creates an ethical dilemma for raters, who are very likely to feel incentives and pressure to inflate the ratings they give. Earlier in this chapter, we suggested some practical methods for reducing the direct supervisor's visibility and responsibility in performance appraisal (e.g., by collecting information from a wide range of sources but feeding it back as a single average). We advanced this suggestion as a way to solve some practical problems, but this approach may also help solve some ethical problems. Performance appraisal, as it is usually implemented, puts managers and supervisors in a position where they will both be tempted and pressured to bend the truth, by giving high ratings even when low ones are deserved. Our first concrete suggestion for the ethical practice of performance appraisal is that organizations should try to avoid placing supervisors, managers, and employees in situations that consistently and predictably pressure them to provide dishonest ratings and feedback.[5]

Second, performance appraisal systems place raters in situations where they have competing moral demands, in particular, doing what is best for the organization versus doing what is best for the individual being evaluated. Suppose that it would be beneficial to the organization to give Sam a low rating, but beneficial to Sam (both because it will get him a raise and because it will help motivate him to improve in the future) to give him a high rating. Which choice should the rater make here? In one sense, the rater is acting as an agent of the organization and is obligated to uphold their interests, but this rater is also an individual with a moral duty to treat others with respect and consideration. We are not sure that there is a clear and simple ethical principle that can be used to resolve these conflicts. The best suggestion we have is for organizations to work to reduce these conflicts. For example, organizations that use performance appraisals mainly to help them identify developmental needs and opportunities probably put raters in fewer conflicts of this sort than organizations that use appraisals to make decisions about promotions, raises, or layoffs.

Finally, managers and supervisors are faced with self-serving motivations to inflate performance ratings. As we have noted in several previous chapters, the key job of a manager is to create conditions that increase the likelihood that his or her subordinates will perform well. A manager who gives low ratings—that is, reporting that his or her subordinates are *not* performing well—is taking a big risk. Our final suggestion is that organizations should take proactive steps to reduce the temptation to make yourself look good by providing inflated ratings. To our knowledge, few organizations reward raters who submit performance ratings that appear to be reliable and valid measures of performance, and even fewer sanction raters whose ratings are obviously inflated. This is another case where the collection of rating information from multiple sources might help. If a supervisor gives ratings of "Greatly Exceeds Expectations" on just about all aspects of performance, but peers, subordinates, and others who do not have self-serving interests give lower evaluations, organizations need to at least consider the possibility that the supervisor is giving biased evaluations. It may turn out that he or she is doing so for good reasons (e.g., on the expectation that high ratings will be more motivating than low ones), but if there are substantial reasons to believe that a supervisor is distorting his or her evaluations, it is at least worth discussing this trend. You can make a case that it is unethical for organizations to ignore evidence that ratings are either accurate or distorted. If an organization wants honest evaluations, it is surely better to structure its HR system in such a way that honesty is recognized and appropriately rewarded.

Summary

Performance appraisal researchers spent much of the period between the 1930s and the 1990s trying to develop ways of improving the reliability and accuracy of performance measures. We believe these efforts led to some useful interventions, but that the more pressing problem is to find ways to make performance appraisals better by making the task easier and less risky for raters. The ultimate goal of performance appraisal is not simply to develop a system that accurately measures performance, but rather to build a system that helps to improve the efficiency and effectiveness of the organization. Ultimately, the only type of appraisal system that is likely to accomplish this goal is one the users trust and believe in.

In our view, trust is the single most important factor in determining the success or failure of performance appraisal systems. If supervisors, managers, and employees do not trust that evaluations will be reasonably accurate and fair, and that the organization will use information from performance appraisals in a consistent and reasonable fashion, it hardly matters what scales you use or what your rater training program looks like. Ultimately, the key to building better performance appraisal is to build better organizations. In particular, if you want performance appraisal to succeed, we believe it is critical that the culture and climate of the organization supports mutual respect and trust among the users of this appraisal system.

Traditional performance appraisal systems often have a set of built-in moral challenges. In that abstract, most people believe that it is important to be honest, but performance appraisal systems predictably put raters, ratees, and human resource managers in positions where there is strong pressure to bend the truth (by inflating ratings, by working to build positive but misleading impressions, by pretending that performance appraisal is important when in fact it is largely ignored). We believe that organizations can and should take a number of steps to reduce the pressure and temptation to distort and misrepresent and that reducing these predictable moral hazards is likely to have real benefits.

In this book, we have devoted a great deal of attention to the challenges and shortcomings performance appraisal systems face. There is no doubt that performance appraisal is hard work, and that many organizations fail to get as much out of their performance appraisal system as they should. We continue to believe, however, that it is possible to develop performance appraisal systems that contribute to the success of organizations, and that organizations that take effective steps to build a climate of mutual trust and respect have a real chance to succeed in the task of evaluating the performance of its members and using that information to make sound decisions. Performance appraisal can work, and the process of making performance appraisal systems effective is one that may provide lasting benefits to organizations. An organization that successfully develops and maintains the climate of mutual trust and respect that allows performance appraisal to succeed will become a better place to work, making it easier to attract and retain talented and dedicated employees. Performance appraisal is a real challenge, but we are more optimistic than ever that organizations can rise to this challenge, and that organizations that succeed in developing a good performance appraisal system will find themselves on the road to broader and more general success.

Case Study: Implementing a Performance Appraisal System

The case study in Chapter 13 describes the development or designing phase of restoring an old, ineffective performance appraisal system (for the nonprofit organization illustrated in the Chapter 13 example). This scenario is a continuation of that example, illustrating how the consultant might implement or deliver the new system.

Now that you have worked with the committee to develop a performance appraisal system that better addresses the organization's needs and aligns with its values, the committee is eager for your guidance in how to execute the new system.

Communication. As it was discussed in Chapter 13 and this chapter, trust is crucial. Involving the employees and managers in the development phase is essential, and it is also necessary to include them in the execution phase to let them know their voices were heard. Communicating thoroughly with the raters and ratees about the changes that will be implemented contributes to the success and transparency of the program (e.g., What are the goals of the new program? What will change for you? Why is it needed? What will happen? What are the benefits and improvements?). This is also a time to set clear expectations so everyone is clear on the goals and objectives from the start. There should be no surprises by the time the new performance appraisal launches.

Training and preparation. Along the same lines as communication, training should occur to build understanding and set expectations for the new system. Raters could benefit from role-play scenarios, calibrate examples with other managers, a tutorial of the new paperless system, specific instruction on goal setting, long-term vision of the performance appraisal, and in many circumstances frame of reference training might be helpful. Further, training on how to lead a face-to-face meeting could help raters feel more comfortable with the process (e.g., discuss active listening, nonverbal body language, what to discuss, how to convey the right tone, how to handle difficult reactions from employees). This could also include protocol for disciplinary action, developmental action, and delivering criticism where necessary. As always, a discussion about legal concerns should be included in the training (e.g., do not use discriminatory language, do not make promises where not completely verified, avoid discussing personality traits or other attitudes/behaviors that are not in the employee's scope to change or develop).

Ongoing feedback and corrective guidance. The rater's job is not always finished after the performance appraisal meeting. When goals are not met or people need help achieving their goals, follow-up should include ongoing feedback, support, and resources.

Program evaluation. Performance appraisal should not be a static exercise. The program should be evaluated, updated, refined, and developed over time to ensure relevancy and validity. This is a complex process because the effectiveness of the program must be evaluated by measuring appropriate criteria (i.e., what was the performance appraisal supposed to achieve?). Further, valuable metrics can be collected afterward, such as data to determine if performance is improving, and establish satisfaction (employees and managers alike) with the new system.

Notes

1. We recognize that even in systems where the direct supervisor is the only one to rate an employee's performance, the judgments of this supervisor are not likely to be absolute, in the sense that there may be some check-off by higher-level management, especially if ratings drive high-stakes decisions.
2. In Chapter 8, Figure 8.1 illustrates this pattern.
3. For example, General Electric has invested in mobile apps to aid supervisors and managers in conducting performance evaluations. See https://qz.com/428813.
4. Six and Sorge (2008) distinguish between espoused values and values that are lived or experienced.
5. There are a number of analogs in the law (doctrine of attractive nuisance, which states that landowners can be held liable if children are injured on their property that contains an object that is known to attract children to trespass and engage in unsafe behaviors, such as a trampoline or an unfenced swimming pool) and in theology (doctrine of occasion of sin—individuals have an obligation to avoid circumstances that they know are likely to incite or entice them to commit immoral acts)