CHAPTER 8

Using Multisource Feedback for Employee Development and Personnel Decisions

Anthony T. Dalessio

"Seek to know yourself, which is the most difficult knowledge to acquire that can be imagined." This quote from the character Don Quixote is an elegant statement of the importance and difficulty of gaining self-knowledge, which has been extolled as a human virtue throughout philosophy and literature. One important key to self-knowledge is understanding how others perceive our behavior. This knowledge can be used to gain insight into the pattern and texture of competencies we possess, and can be used to help us leverage our strengths as well as identify areas for self-development and self-acceptance. Multisource feedback can be used as one method for obtaining information about how others perceive us and how we perceive ourselves.

Multisource feedback is defined as evaluations gathered about a target participant from two or more rating sources, including self, supervisor, peers, direct reports, internal customers, external customers, and vendors or suppliers. By definition, what has been referred to as 360-degree feedback in the literature falls under the heading of multisource feedback, and includes an opportunity to receive feedback from individuals in positions above, below, and lateral in the organizational hierarchy, as well as from individuals outside the organization. Feedback processes referred to as upward feedback and team member feedback (or peer feedback) also fall under the rubric of multisource feedback when evaluations are provided by additional sources such as self-ratings.

Brutus, Fleenor, and London (1998) have likened the multisource feedback process to a thermostat-control system that begins with participants' setting performance goals on job-relevant competencies, then gathering feedback on competencies from various sources, effectively processing the feedback, making adjustments by changing behavior based on the feedback, and finally looping back to establish new performance goals. In applications of multisource feedback, evaluations of a participant are typically gathered from the sources through ratings and written verbatim comments on a set of common items assessing job-relevant dimensions or competencies. This information is then compiled into a feedback report for the participant that protects the anonymity of the individual raters. The participant is provided with instruction on how to read and interpret the feedback report, as well as how to use the information to construct developmental plans and follow-up actions.

Based on a recent telephone survey of organizations and consulting firms, London and Smither (1995) concluded that multisource feedback currently is widely used by businesses. For example, seven of the consulting firms contacted in the survey indicated that they each have applied multisource feedback with over a hundred companies. Three of the consulting firms indicated that they have sold their multisource feedback instruments to over a thousand organizations. Companies are also spending large amounts from their training and development budgets on multisource feedback. Romano (1994) noted that Personnel Decisions, Inc., estimated that companies spent $152 million on multisource feedback processes in 1992. London and Smither point out that multisource feedback has also received attention in popular practitioner publications. Another indication of the popularity of multisource feedback is the recent establishment of the Upward Feedback Forum (UFF) by practitioners in organizations with active multisource feedback processes in place. The UFF is an informal consortium of approximately twenty-five

Note: The author thanks Dick Reilly, Jim Smither, Amy Hirsch, Lisa Kuller, John Mallozzi, and Anna Marie Valerio for their helpful comments.

major organizations that began meeting semiannually starting in mid-1993 to discuss best practices regarding multisource feedback (Timmreck & Bracken, 1996). Also, multisource feedback has become increasingly popular in industrial and organizational (I/O) psychology. For example, at the 1995 through 1997 Society for Industrial/Organizational Psychology conferences, a total of fourteen symposia, practitioner forums, or workshops were devoted to the topic.

What has spurred the recent popularity of multisource feedback programs? Reasons include the perennial issue of dissatisfaction with traditional performance appraisal, and an increasing focus by organizations on issues of empowerment, participation, and customer focus (Bracken, 1996). Multisource feedback is one way to open the appraisal and feedback process and include the perspectives of internal and external customers, as well as others in the organization who work with the employee, besides just the supervisor. The process empowers the employee to use the feedback to achieve growth and self-development. The popularity of the technique also stems from its use as a way to communicate critical organizational behaviors and values that need to be practiced to move the organization forward (London & Beatty, 1993). It can be one component of a larger culture change effort to convey the importance of critical organizational competencies and values. Finally, a multisource feedback process may be a more efficient, less time-consuming method for providing participants with developmental feedback than other methods such as traditional assessment, which many organizations no longer have the resources to support.

With the increasing popularity of multisource feedback processes, many practitioners will be faced with a number of important implementation questions. Research and practice relevant to these questions will be examined in this chapter, and recommendations for practice offered. After a discussion of applied measurement issues associated with multisource feedback processes and the practical implications of current research, the following implementation questions will be addressed: What issues need to be considered and resolved when planning a multisource feedback process, including use of the process for development versus for personnel decisions, communications to participants and raters, and working with consultants? How should a multisource feedback questionnaire be constructed, including issues of identifying and developing competencies, item development, response scales, and open-ended questions? What issues need to be resolved regarding administration, such as choosing raters and developing rater training? How should feedback reports be constructed, including data aggregation and presentation? How can an organization facilitate support for developmental activities, including the feedback workshop and follow-up activities? How can it evaluate and revise the process?

Measurement and Validation: The Source of Multisource Feedback

Often practitioners suggest that multisource feedback instruments need not meet strong validation standards, if the instruments are used only for developmental feedback purposes rather than for personnel decisions. Nonetheless, regardless of the purpose, a multisource feedback instrument must still meet the same standards of validity (Bracken, 1996; Dalessio, 1996). Even when the instrument is used purely for the purpose of developmental feedback, the participant will still need to be provided with a valid measure of job-related competencies for constructing a job-relevant developmental plan. And if the results of the instrument are used either intentionally or inadvertently to make a job performance, compensation, promotion, or downsizing decision, it must meet the validation standard for any instrument used for a personnel decision.

Tornow (1993) and Dunnette (1993) provide an interesting dialectic for framing a discussion of measurement issues for multisource feedback instruments. Tornow contrasts a traditional psychometric measurement approach for multisource feedback instruments with a practitioner's approach. On the traditional psychometric end of the continuum, Tornow views the goal of the scientist as "to reduce the error of measurement, and rater variations, in order to enhance the accuracy of measurement" (p. 222). On the other side of the continuum, according to Tornow, "the manager/practitioner values multiple perspectives of different raters because they represent significant and meaningful sources of variation from which much can be learned" (p. 222). Dunnette elaborates on these points by stating that "good measurement—no matter how elegant or wonderful—is sterile unless it is accompanied by meaningful findings. Likewise, provocative interpretations may be rendered illusory if they cannot be firmly supported by solid methodology" (p. 378). For scientist-practitioners, what methodologies constitute evidence for reliability and validity of multisource feedback instruments, yet still allow for provocative interpretations that practitioners can provide to participants to assist in meaningful development and growth?

Reliability

Reliability can be defined "as the extent to which a set of measurements is free from random-error variance" (Guion, 1965, p. 30). Multisource feedback instruments should be able to adequately demonstrate internal consistency reliability for the dimensions assessed with coefficient α reliabilities in the .70 range (Van Velsor & Leslie, 1991a), although lower coefficients may be acceptable, depending on the construct being assessed. Also, reasonably systematic agreement among raters *within* a source (for example, peers) would be expected for multisource feedback. A meta-analysis by Conway and Huffcutt (1996) concluded that raters who are supervisors, peers, and subordinates can provide a reliable *within*-source index, and that within-source reliabilities for peers and subordinates can be increased by including from four to six raters for each of these sources. Conway and Huffcutt report lower within-source agreement for subordinates (.27) than for other sources (supervisors, .51; peers, .39). These results may be partially due to supervisors' behaving differently toward different subordinates, as suggested by the vertical dyad linkage model of leadership (Dansereau, Graen, & Haga, 1975).

Lower agreement *between* rating sources (for example, supervisors and peers, or peers and subordinates) may not necessarily be random error variance, because different rating sources may perceive real behavior differences by the participant (Conway & Huffcutt, 1996; Tornow, 1993). Also low test-retest correlations may not provide evidence for poor reliability, especially when retest intervals are long enough to allow participants to develop and change, which is the major objective of most multisource feedback processes. So reliability for multisource instruments may be best estimated through internal consistency methodology and interrater agreement methodology within sources.

Validity

Validity is traditionally defined as the extent to which an instrument measures what it is purported to measure (Guion, 1980). Although the concept of validity can generally be unified under this theme, evaluative judgments regarding evidence for validity can be drawn from all three conventional aspects of validity—content, criterion-related, and construct (Guion, 1980). What types of evidence would result in judgments regarding evidence for validity of a multisource instrument?

Content Validity

Content validity, or a content-oriented strategy, is an evaluation of the extent to which the items in the multisource feedback instrument represent the defined job performance content domain (American Educational Research Association, American Psychological Association, & National Council on Measurement Education [AERA et al.], 1985; Goldstein, Zedeck, & Schneider, 1993; Guion, 1980; Society for Industrial and Organizational Psychology [SIOP], 1987). Evidence for the content validity of a multisource feedback instrument can be established through a traditional job analysis–content validation process (Goldstein et al., 1993) where subject matter experts (SMEs)—people who know the job—are used to identify important job tasks; define and identify the important knowledge, skills, and abilities required to perform the tasks; and then link multisource feedback items to the critical knowledge, skills, and abilities. The process can be accomplished through facilitated SME meetings and structured questionnaires (Goldstein et al., 1993).

Most applications of a multisource feedback process involve developing a content-valid questionnaire that can be used with managers at many different levels, for example, senior management through first-level management, and with managers in many different functional areas, for example, marketing, sales, production, operations, human resources, finance, and so on. For such applications, a content-validation process would involve using SME meetings and questionnaires to behaviorally define generic management competencies and establish their importance for the successful performance of all the relevant management jobs in the organization. Behavioral items on which employees could initiate change would need to be developed and then linked to the management competencies.

Criterion-Related Validity

Evidence for *criterion-related validity* is based on the extent to which evaluations from various sources are related to outcome criteria (AERA et al., 1985; SIOP, 1987). Two major issues that need to be taken into account when establishing criterion-related validity are first, that the predictor consists of ratings from multiple sources, and second, that the outcome criterion needs to be a measure of effective managerial job performance. For criterion-related validity, each particular source or combination of sources, for example, peers and customers, should be considered separately. It should be kept in mind that ratings from some sources may not necessarily need to demonstrate criterion-related validity to provide valuable information. Ratings from each source represent a legitimate view of job performance, if the source of variation is providing ratings on behaviors linked to important organizational competencies and is used as input to a comprehensive job-related developmental plan.

The other side of the equation is the outcome criterion measure. Choosing an appropriate criterion should be based on clearly defining effective managerial performance (Van Velsor & Leslie, 1991a). Effective managerial performance may be defined as profitability of the unit, high team performance, high job satisfaction for the unit, or low unit turnover. Ratings from particular sources may be more logically related to certain criteria. For example, subordinate ratings may be related to job satisfaction and turnover for the unit, whereas supervisor ratings may be less related to these criteria and more highly related to unit profitability. Also, outcome criteria may not be practically available for some sources, for example, self-ratings. The outcome criterion used to establish criterion-related validity should be under the control of the manager, and will be less appropriate for use in establishing criterion-related validity to the extent that it is influenced by variables other than effective management behaviors (Van Velsor & Leslie, 1991a).

Finally, evidence for criterion-related validity can also be established by demonstrating that the instrument is able to differentiate between groups of managers considered overall to be

performing effectively on a relevant criterion variable compared to those considered to be performing ineffectively.

Construct Validity

Evidence for *construct validity* focuses primarily on the extent to which scores on the instrument measure the psychological characteristics of interest (AERA et al., 1985; SIOP, 1987). The process of construct validity involves establishing meaningful, logical patterns of relationships among items, scales, and sources of measurement. The value of multisource feedback is that information about the participant does not come from just one perspective, such as the supervisor, but rather from many different perspectives. The assumption is that each source provides relevant information about a performance dimension, but may evaluate the participant somewhat differently because the participant actually behaves somewhat differently toward each source. Variations also arise because each source does not have the opportunity to view the same behaviors on the dimension, or because the sources differ somewhat in terms of which behaviors they view as most critical for a dimension. This assumption may lower patterns of relationships between certain sources of measurement, for example, subordinates and peers, and subordinates and supervisor, as compared to other sources, such as supervisors and peers (Conway & Huffcutt, 1996). However, these results would not indicate that subordinates' ratings, for example, are useless or not valuable. Subordinates' ratings still provide a legitimate source of feedback to the participant that needs to be considered and used for planning developmental activities.

The constructs measured by the multisource instrument may not be defined by precisely the same behaviors for each source, because the sources differ in their opportunities to view the participant's behavior. Most multisource feedback processes allow raters to endorse an option of "not applicable" if they have not had a chance to view a behavior adequately. The most obvious example involves items that ask about supervisory style or about performance feedback. These items are observed clearly by direct reports but much less clearly by sources such as customers. Factor structures for multisource feedback items may be difficult to establish with data from all sources combined, because a certain percentage of raters from a given source may not be able to evaluate some behavioral items. Within-rating-source data may produce clearer factor structures.

Finally, evidence for construct validity can be established through correlations between ratings made by a particular source and other measures of the same competency, such as assessment center ratings or certification tests.

Aspects of Validation: Summary

Given the nature and objectives of multisource feedback, the key aspect for inferring validity for an instrument is establishing that the items represent the content of the job performance domain, that is, demonstrate genuine content validity. Criterion-related validity for the multisource feedback process can be established when ratings predict appropriate outcome criteria, but not all sources necessarily need to be directly related to outcome criteria to demonstrate value as part of the process. Although ratings among some sources would be hypothesized to show convergence, other sources may not show strong convergence yet still provide meaningful information to participants. Constructs may be defined by somewhat different behavioral items for different sources. Appropriate measures of the same construct should be related to multisource feedback ratings.

Multisource Feedback Research: Implications for Practice

The results of research conducted on multisource feedback processes provide suggestions and guidance that practitioners should consider for implementation purposes. Nearly all multisource feedback research has been completed in the context of a developmental setting rather than a personnel decision setting, so results primarily apply to developmental uses. Research falls into three main categories: implications for administration, implications for feedback of results, and implications for outcomes of the process.

Implications for Multisource Feedback Administration

Should Raters' Evaluations Be Anonymous or Identified?

In an experimental field study of upward feedback, Antonioni (1994) randomly assigned thirty-eight managers and their subordinates from an insurance company to either an anonymity or an accountability condition. In the accountability condition, raters signed their completed upward feedback forms and supervisors received completed individual subordinate questionnaires. In the anonymity condition, by contrast, raters did not sign and supervisors received reports only of summary results. Participants and raters were informed that the process was being used solely for the purposes of developmental feedback, and would not be used for personnel decisions. The results indicated that ratings were less inflated in the anonymous condition, and subordinates were more comfortable providing ratings in this condition. The research suggests that better quality ratings will be provided to participants when raters are anonymous. This conclusion is also consistent with survey data of raters' reactions to an upward feedback process reported by London and Wohlers (1991). When subordinates were asked, "Would you have rated your boss any differently if feedback had not been given to them anonymously," 24 percent indicated that they would have done so (p. 388). Supervisors in the Antonioni study did view the process more favorably when subordinates were identified, which may improve their acceptance of the results, but the results still suggest the best practice is rater anonymity. Future research should investigate the anonymity variable for potentially different effects with other rating sources, including peers, supervisors, and internal and external customers.

Is the Quality of Ratings Affected by Use for Developmental Versus Appraisal Purposes?

Research has indicated that ratings collected for the purposes of personnel decisions are significantly higher than ratings of the same individual collected for development or feedback purposes. For example, Farh, Cannella, and Bedeian (1991) have demonstrated that peer ratings from class project teams of university management students showed greater leniency and halo and were less differentiating when gathered for evaluation purposes than when gathered for developmental purposes. Other studies also "have shown that ratings collected for administrative purposes (for example, salary administration) are significantly higher than ratings of the same individual collected for other purposes, such as feedback or research" (Murphy & Cleveland, 1995, p. 246). Furthermore, London and Wohlers (1991) report that when employees who participated in an upward feedback process were asked whether they would have rated their boss differently if the feedback was used for performance appraisal (rather than development), 34 percent indicated that they would have done so.

The research suggests that when multisource feedback is used for personnel decisions rather than purely developmental purposes, practitioners should be prepared for ratings that are potentially contaminated by rating errors, which may make distinguishing between genuinely good versus mediocre performers difficult and limit the value of such ratings to guide development. Attempts could be made to reduce these errors through more intensive rater training and more extensive employee communications about the importance of providing accurate ratings.

Who Should Choose the Raters?

Research is needed to determine how the quality and acceptance of ratings is affected by participants choosing their raters versus others providing input. No research has been conducted on this question to date.

Do the Rating Sources Included in the Multisource Feedback Process Affect Participants' Attitudes Toward the Process?

Bernardin, Dahmus, and Redmon (1993) examined whether the attitudes of first-line supervisors in a large federal agency differed toward upward appraisal depending on whether they received feedback from both their manager and their subordinates, from their manager only, or from their subordinates only. The results indicated that the most positive attitudes were displayed toward upward feedback when supervisors received feedback from both their manager and their subordinates. Supervisors who received feedback from just their subordinates expressed more cautious attitudes toward upward appraisal, as well as more concern about possible negative implications such as supervisors focusing too much attention on simply pleasing subordinates. Bernardin et al. (1993) suggest that the implication of their results for practice is that, whenever possible, a more comprehensive approach to feedback should be implemented where feedback from all qualified sources is sought. With a more complete multisource feedback process, attitudes toward subordinate feedback as well as feedback from other sources may be more tempered.

Implications for Feedback of Multisource Results

Should Feedback to Participants Focus on Discrepancies Between Self-Ratings and Others' Ratings?

Research that has focused on examining discrepancies between self- and others' ratings provides insight for interpreting multisource feedback results. The research results suggest implications for the meaning of particular discrepancies and suggest prescriptive actions that participants should consider as a result of their feedback.

Meta-analytic work by Harris and Schaubroeck (1988) has concluded that ratings between peers and supervisors agree to a greater extent than ratings between self and others (that is, self and peers; self and supervisors). Self-ratings also tend to be higher on average than ratings provided by others (for example, Bass & Yammarino, 1991; Brutus, Fleenor, & McCauley, 1996; Harris & Schaubroeck, 1988; Smither et al., 1995). Nilsen and Campbell (1993) have shown that the size of these self-other discrepancies tends to show a high degree of stability over a period of time using scales from the Campbell Leadership Index. The median correlation reported by these researchers between self-other discrepancies over a period of one month for a sample of thirty-one female managers was .71. Although the time interval was short in this study, it seems appropriate, since much longer intervals could have resulted in changes due to systematic intervening variables such as experience or training. Nilsen and Campbell also provided data showing that self-other discrepancies were stable across different types of measures, such as personality and skill measures. Without interventions, participants are likely to display stability in these discrepancies over a period of time and across measures.

The size and nature of these discrepancies may be related to managerial effectiveness. For example, Bass and Yammarino (1991) conducted a study where 155 Navy officers completed self-ratings of leadership behavior, and their subordinates also rated the leadership behavior of these officers. Also, Navy records for the 155 officers were obtained to gather performance and promotion data used as an index of each officer's success. The results of the study indicated that generally less successful officers showed greater discrepancies between self- and subordinate's ratings of leadership behavior, whereas more successful officers showed lower discrepancies. In a study addressing a similar question, Van Velsor, Taylor, and Leslie (1993) analyzed data from three samples of managers who completed self-ratings using Benchmarks, which is a multisource feedback instrument developed by the Center for Creative Leadership. The samples consisted of 648 managers randomly selected from the Benchmarks database, 168 upper-level managers from a Fortune 100 organization, and 79 hospital administrators. Subordinates also rated these managers using Benchmarks. The managers were coded into three groups—underraters, agreement, and overraters—based on the discrepancy between average self-ratings and average subordinate ratings across fifteen of the sixteen Benchmark scales. (The self-awareness scale from Benchmarks was not included as part of the coding process.) Overraters in all three samples received the lowest mean overall rating from subordinates across all the Benchmark scales, and had the highest self-ratings. This is consistent with the Bass and Yammarino (1991) results. However, in the Van Velsor et al. (1993) study, the underrater group (rather than the agreement group) received the highest mean overall rating from subordinates, and the lowest self-ratings. Analysis of the self-awareness scale from Benchmarks revealed that subordinates evaluated underraters as highest on self-awareness.

Workshop facilitators should be aware of and use this research to underscore the need for managers to be cognizant of, and consider, others' perceptions of their behavior. This information is especially important if a participant has consistently high self-ratings on the feedback report. The research suggests that these managers may be evaluated as less effective by superiors and direct reports, as well as be perceived as being less self-aware of their behavior. As part of their developmental planning efforts, these managers in particular may need to focus on bringing their self-views more in line with others' views.

Practitioners should also be aware that recent research by Fleenor, McCauley, and Brutus (1996) suggests that level of performance needs to be considered as much as self-other discrepancies in interpreting results and focusing developmental efforts. Facilitators need to focus participants' attention not only on self-other discrepancies but also on behaviors receiving low to average ratings as opportunities for development, and behaviors with high ratings as strengths to be leveraged.

Subordinate-other discrepancies in ratings also received attention in the research literature recently. A meta-analysis by Conway and Huffcutt (1996) confirmed the Harris and Schaubroeck (1988) conclusions and extended this research to show that agreement between peers and supervisors is also higher than agreement between subordinates and peers and subordinates and supervisors. These results suggest that it is important to include subordinates' ratings in multisource feedback, because of their unique perspective relative to peers and supervisors.

Should Participants Be Open to Seeking Negative Feedback?

One approach that could be suggested to participants in multisource feedback workshops is to be open to seeking negative feedback about their behavior from supervisors, subordinates, and peers, as a way to gather information that can be used to narrow the gap between self- and other perceptions, and to improve in developmental need areas. Based on questionnaire data from a sample of 387 mid-level managers in a public service agency, Ashford and Tsui (1991) concluded that the tendency to seek negative feedback was positively associated with self-other agreement for supervisors, subordinates, and peers, as well as with overall effectiveness ratings from all three groups. Their research also suggested that the tendency to seek positive feedback was not a useful strategy for the managers. Results indicted that seeking positive feedback was negatively related to effectiveness ratings from all three sources, and with self-other agreement for subordinates. Positive feedback seeking was not related to self-other agreement for supervisors and peers.

Implications for Outcomes of Multisource Feedback Processes

Does Instituting a Multisource Feedback Process Improve Performance?

This question will eventually be asked by senior management at most organizations. Research on this topic has indicated that performance improvements do result from multisource feedback processes, although improvements appear to be moderate. Practitioners can cite this research as support for instituting or continuing a multisource feedback program within an organization. In a study by Hazucha, Hezlett, and Schneider (1993), forty-eight managers from a large

Midwestern utility company, and their supervisors, peers, and subordinates, completed the Management Skills Profile (MSP) (which assesses nineteen job-related dimensions) at two points in time, two years apart. The results showed that average ratings by others increased from Time 1 to Time 2. Also the correlation between self-other ratings were more similar at Time 2 ($r = .32$) compared to Time 1 ($r = .07$).

Research conducted by Smither et al. (1995) with data from an upward feedback program showed similar results when effects for regression to the mean were controlled. In the Smither et al. study, data were collected from 238 managers and their subordinates in the international operations division of a large organization at two points in time, about six months apart. Ratings were gathered on thirty-three behavioral items. Analyses of subordinates' ratings showed that the managers whose initial level of performance was moderate or low improved over the six-month period. The improvement could not be attributed solely to regression to the mean. Reilly, Smither, and Vasilopoulos (1996) extended the Smither et al. study by following 171 of the original 238 managers for a third administration of the upward feedback questionnaire, and 92 of the managers for a fourth administration two and a half years later. Managers whose initial level of performance was low sustained improvements over the later administrations. Atwater, Roush, and Fischthal (1995) obtained similar results, indicating that leaders' behaviors improved after receiving feedback from followers in a sample of 978 student leaders at the U.S. Naval Academy. A recent study by Walker (1997), examining upward feedback for 252 managers in a financial institution, also showed steady improvements in subordinates' ratings of their supervisor on a composite of leadership behavior items over five yearly administrations. Consistent with previous research, Walker also found improvements primarily for supervisors with initial low and moderate ratings, and these improvements could not be accounted for solely by regression to the mean. An additional important finding from Walker's study was that managers who held feedback sessions with subordinates showed greatest improvements in ratings. So holding feedback meetings may be one potential key to improvement. Walker also showed that subordinates' ratings were correlated with independent annual performance appraisal ratings made by supervisors with stronger correlations in later administrations.

Bernardin, Hagan, Ross, and Kane (1995) examined not only the effects of a multisource feedback process on managerial performance ratings but also outcome variables such as sales volume and turnover with a sample of forty-eight assistant store managers in a retail clothing chain. The managers were divided into two groups, one of which received multisource feedback while the other served as a control group and received only top-down appraisal. Data were collected from each of the groups at two points in time, one year apart. In the multisource feedback group, ratings were gathered on twenty behavioral items from the participant, subordinates, peers, and supervisor at Time 1 and Time 2. Also "professional customer" role-players came into the stores at Time 1 and Time 2 to test system policy as implemented by each manager, and rated the manager using a fifteen-item instrument. In the control group, supervisor and self-ratings were gathered at both times, and customer ratings at Time 2 only. Also, two outcome measures (voluntary termination of employees supervised by the assistant manager, and "sale to target") were collected for both groups at both times. Mean subordinate, peer, and customer ratings increased significantly at Time 2 for the multisource feedback group. Ratings by supervisors and self-appraisals did not improve significantly for either group. The two groups did not show significant differences at Time 1 or Time 2 on either of the outcome measures. The results show moderate support for the effects of multisource feedback on performance variables. Further research on the relationship between multisource feedback and outcome variables is needed, although at least one rating source included in multisource feedback (that is, peer assessments) has shown relationships with outcome variables (for example, see Kane and Lawler, 1978; Mayfield, 1970, 1972).

Do Managers Need to Receive Feedback to Show Improvements in Performance?

Or do managers improve performance as a result of just participating in a multisource feedback process? In the field studies conducted by Smither et al. (1995) and Reilly et al. (1996), managers with initially low or moderate performance improved whether or not they received a written feedback report. Simply participating in the multisource feedback process may have familiarized the managers with the organizational competencies and helped them focus on improvement. These results have also been demonstrated in a more controlled experimental study conducted by Dominick, Reilly, and McGourty (1997). The participants in this study were seventy-five graduate and undergraduate students in an organizational behavior and group dynamics course who participated in the study as part of a class assignment. The participants were randomly assigned to teams of four or five, and then the teams were randomly assigned to one of three conditions: feedback, exposure, or control. In the feedback condition, participants completed ratings on themselves and teammates following completion of a first group task. Before a second group task, they received their written feedback reports and were given time to review the feedback. Participants in the exposure condition also completed ratings on themselves and teammates following the first task, and were told they would receive feedback sometime in the future, but did not actually receive written feedback before the second group task. Participants in the control condition completed a questionnaire on task content, but did not complete self- or team member ratings. Videotapes of each team's second task were made and then evaluated by experienced assessors on the same rating form completed by participants in the feedback and exposure conditions. The results showed that assessor ratings were significantly higher both for the feedback versus control group and for the exposure versus control group, but were not different between the exposure and feedback group. Taken together, these three studies suggest that instituting a multisource feedback process can help communicate and reinforce key organizational competencies and culture change initiatives. Also, managers who do not receive written feedback in upward feedback programs (because less than three subordinates rated them) may still benefit from the process (Smither et al., 1995).

Summary of Research Implications

Based on this research, there are several major implications for practice. First, participants in multisource feedback processes are more likely to receive higher quality ratings if ratings are anonymous and used just for developmental feedback. Including input from all sources is also an important planning consideration. Second, workshop facilitators should emphasize the importance of developmental planning that considers bringing self-other ratings more in line, especially when self-ratings show a tendency to be consistently higher than others' ratings. Facilitators also should encourage participants to consider the level of feedback received on competencies in developmental planning. Also, encouraging participants to seek more negative feedback, if this is not being done, may be helpful for improvement. Finally, participation in multisource feedback processes appears to be related to subsequent improvements in performance. Greater improvements in performance may be associated with holding feedback sessions with subordinates.

Implementation of Multisource Feedback Processes for Best Practice

When practitioners develop and implement a multisource feedback process, there are a number of questions that need to be addressed at each step of the process. Possible options exist for completing each step and resolving the questions. These options will be discussed along with recommendations for best practice based on research and practitioner experience.

Planning the Multisource Feedback Process

How to Link the Multisource Feedback Process to Important Organizational Objectives?

One of the best methods for linking a multisource feedback process to organizational and human resource strategies is to base the process on competencies that have been identified as key to the organization's success now and in the future. The multisource feedback process can be used both as a mechanism for communicating these competencies and as a way to provide feedback and develop these competencies. To the extent that these competencies are the basis of other human resource systems, such as performance appraisal, compensation, or succession planning, the multisource feedback process will also provide relevant feedback and input to these systems.

The multisource feedback process can also be linked to organizational strategies by integrating it with organizational change efforts. For example, if the organization is moving toward a team-based environment, the multisource feedback process can be integrated into this change effort by providing feedback on team skills as part of an organization-wide development program.

How to Gain and Display Senior Management Support for the Process?

Linking the process to important organizational strategies is the first step to gaining senior management buy-in to the multisource feedback process. Besides these links, senior management will also need to be convinced that the process is resulting in improved performance. This consideration may not be critical at the initiation of the process, but it certainly will be important to sustain the process. Establishing that the process is related to improved performance can be done through citing the previously discussed research that has shown positive improvements. Such research can also be planned as part of the process, although collecting and analyzing the data can become quite complex. (For examples of such research, see Reilly et al., 1996, and Smither et al., 1995.) The value of the process to senior management can also be demonstrated by defining the outcomes in measurable terms, such as having all the participants write a developmental plan and incorporating specific activities from the plan into their performance appraisal objectives. These measures can be monitored and summarized.

Once senior management buy-in is obtained, it is also critical to the success of the process to demonstrate management's support to participants and raters. In large organizations, cascading the implementation of the process from the senior levels down is a way to demonstrate the importance and support for the process. Also, communication and feedback of results by senior managers to their direct reports for the purpose of action planning and follow-up activities can build support and credibility for the process.

Will the Multisource Feedback Be Used for Development or Personnel Decisions?

This is a major question that needs to be answered as part of the planning process, because it will be the first question that participants ask. Many organizations seem to demonstrate approach-avoidance behavior when deciding whether to use the process for development or for personnel decisions. Dalton (1996) suggests that over the past ten years the use of multisource feedback instruments has evolved from a purely confidential developmental process to include more applications as an appraisal and administrative process. For example, in a survey of twenty-nine multisource feedback

processes conduced through the UFF, Timmreck and Bracken (1995) reported that 28 percent used the process as input to appraisal and 45 percent were moving toward using it as part of the performance appraisal process, although most were still used for development and coaching (93 percent). Yet, in follow-up work, Timmreck and Bracken (1996) suggest that about half the companies either using or intending to use multisource feedback for appraisal purposes backed off from this use. Reasons tend to focus on employee reactions and inflated ratings. Timmreck and Bracken (1996) caution, however, that these conclusions are based on small sample sizes.

Results of a carefully developed, well-validated multisource feedback instrument can either be provided confidentially to the participant, purely for development, or shared with the supervisor as input to personnel decisions. Locating the use of the multisource feedback process at either end of this continuum has benefits as well as costs that need to be weighed by practitioners. If the process is used by the supervisor for personnel decisions, these decisions will benefit by drawing on many sources for input resulting in a more complete picture of the participant's job performance (Bracken, 1996; Murphy & Cleveland, 1995). As discussed earlier, the cost is that the quality of the ratings from these additional sources may be poorer when the raters know their evaluations will be used for personnel decisions (for example, Farh et al., 1991; London & Wohlers, 1991; Murphy & Cleveland, 1995).

When multisource feedback processes are located at the other end of the continuum, and are provided confidentially to participants solely for developmental purposes, better-quality ratings with less leniency may result. These ratings will provide better information to the participant for constructing a development plan. However, if the feedback is completely confidential and left entirely up to the participant, accountability for constructing the developmental plan and completing developmental activities becomes questionable. Participants may have very good intentions in group or individual feedback sessions, but not have the motivation to thoroughly review their results and follow through on developmental activities.

Organizations often begin at the developmental end of the continuum, using the multisource process purely to provide feedback, where only the participant receives the results. With subsequent administrations, the organization may require that the feedback move down the continuum and become more available to the supervisor—first requiring the participant to share the developmental plan with the supervisor; then requiring the participant to share the feedback with the supervisor for the purpose of joint developmental planning; and finally requiring the participant to share the results with the supervisor as input to personnel decisions (that is, performance appraisal, compensation, promotion, or downsizing).

For example, one organization introduced a multisource feedback process such that feedback was available to participants at the same time that midyear appraisals were being completed and developmental plans were being designed between supervisors and direct reports. This administration of the process was purely developmental, where only participants received feedback reports. Participants were told that a second administration of the process would be completed six months later to coincide with the year-end appraisal process, at which time supervisors would also be provided with their feedback reports and use the results as input into the performance appraisal process. The process allowed participants to first receive feedback about strengths and developmental needs, and voluntarily share the information with their supervisor at midyear for the purpose of developmental planning if so desired. The process also incorporated later feedback into the appraisal process.

Rather than move completely to the personnel decision end of the continuum, organizations can adopt various middle positions on the continuum to enhance accountability yet still produce high-quality evaluations for participants. One approach may be to provide the actual multisource feedback report just to the participant in a confidential manner such that it would not be shared with the supervisor, then require the participant to use the report to construct a development plan that would be shared with the supervisor. Agreed-upon self-development activities would then become part of the employee's yearly objectives, and completion of the developmental objectives could be evaluated. Although supervisors do not view the participant's feedback report with this process, they do have access to some information from the process. Sharing information at this level with supervisors may not be inordinately difficult for employees in most instances.

Another middle position on the continuum that can enhance accountability was recently put into practice at a large company. With this multisource feedback process, copies of the feedback report go both to the manager being reviewed and to a consultant (either internal or external). The consultant and the manager meet to review the report, for example, clarify, interpret, identify some areas of strength and areas of concern, and discuss some possible developmental plans and actions. The manager sets up a meeting with the team of direct reports where both the manager and the consultant are present. The manager shares some of the feedback results and tells the group some areas where additional input and clarification would be appreciated. The manager then leaves the meeting, and the group and consultant discuss specific reactions and responses to the manager's questions. The manager returns to the meeting and the group (facilitated by the consultant where necessary) provides more detailed feedback and reactions. The manager summarizes the feedback to make sure that it is accurately understood, and commits to some specific areas and steps for improvement. After the meeting, the manager prepares a developmental action plan and shares it with his or her supervisor.

These positions on the continuum require that at least some feedback be shared with the supervisor. Dalton (1996) has suggested, however, that requiring that such information be shared with the supervisor creates a stressful and threatening situation that may lead the participants to use coping strategies that produce movement away from personal development. Dalton has pointed out that counseling principles suggest that feedback will be most readily accepted in a confidential interaction between a qualified feedback giver and the participant, thus locating the process closely to the purely developmental end of the continuum. The feedback giver, for example, could be an outside consultant who works with the participant as a coach in planning and completing developmental activities. Using outside coaches for large-scale multisource feedback applications may be extremely costly. Another less costly option may be to use consultants inside the organization or peers as coaches, although quality and consistency of the coaching as well as confidentiality would need to be addressed.

There is no definitive answer regarding where multisource feedback processes should be located on the developmental–personnel decision continuum. Practitioners and their clients need to weigh the costs and benefits of any approach in light of business needs. As a general recommendation, multisource feedback processes are better used as developmental tools. Under these circumstances, the process is more easily introduced to an organization in a positive light and is better accepted by participants, but it is still necessary to address accountability for development.

What Up-Front Communication Should Be Provided?

Up-front communication is very important to gain buy-in, and to promote understanding about the purpose, benefits, and expected outcomes of the process, as well as to explain to participants and raters what they will be asked to do, and in what time frame. An essential part of the communication is a clear statement regarding whether the results will be provided solely to the participant for developmental purposes, or will be shared with the supervisor for administrative purposes. The reasoning behind this decision also needs to be communicated. Whether or not results are provided to supervisors in any form, raters and participants should be informed that raters' evaluations will remain anonymous by reporting data for a source only if a minimum number of raters (for example, three or more) provide ratings. If an outside consultant is aggregating the ratings and producing the reports, this procedure can also be communicated as an additional way that rater anonymity will be maintained.

Besides these issues, participants and raters should be provided with clear information that reflects sound planning on both scheduling and content issues relevant to questionnaire development, administration, scoring, and feedback; how frequently feedback will be provided; what follow-up support will be in place once feedback is delivered; and how long the process will be in place. Meetings with participants, the project manager, and the upper-level manager responsible for the project can be held to communicate the information. A brochure containing the information and contact persons also can be provided.

How to Work with an Outside Consultant to Develop and Implement Multisource Feedback?

Many organizations planning a multisource feedback process work with a consultant in the development and implementation. Several models exist for contracting with a consultant that range from using

the consultant's off-the-shelf questionnaire to customizing this type of questionnaire to contracting only for administrative services. To a large extent, the model that is most appropriate depends on the resources and staff that are available within the organization.

The competencies assessed with off-the-shelf products are applicable for most organizations. However, a content validation process linking the competencies to job activities should be completed to provide evidence for the instrument's job-relatedness. The extent to which these competencies adequately describe specific behaviors that can be influenced and changed should also be evaluated. The psychometric properties and national norm database of these products are potentially strong because of the large amount of available data. Van Velsor and Leslie (1991a, 1991b) present a useful framework for evaluating multisource feedback products in terms of psychometric properties, as well as feedback displays and support materials. Van Velsor and Leslie (1991b) also have reviewed a number of multisource feedback products available from vendors. An additional product developed more recently has been reviewed by Dalessio (1996). Finally, the City of Hampton, Virginia, also has available a brief report completed in 1994 that describes several multisource processes and their costs. These reviews will be useful to organizations considering implementation of multisource feedback.

Customization of at least some items on vendor-developed questionnaires may be necessary for the purposes of focusing the competencies or making the questionnaire more congruent with the organization's culture. A balance needs to be struck when modifying items—useful as it is, this process will undercut the advantages of the national norm and psychometric data because the national data will not be applicable to the modified items.

A useful addition to the multisource feedback questionnaires offered by vendors is a technical manual that includes data on the psychometric properties of the questionnaire items (Dalessio, 1996). This information would be helpful to users with technical backgrounds for choosing items to include in a multisource feedback instrument. Such a manual could provide information on the internal consistency reliability of items composing the dimensions, item-total correlations for items within dimensions, the variability of ratings on each of the items for the sources, and average ratings on the items and the dimensions for the sources.

Consultant services could also be used for purely administrative purposes, including distributing, receiving, entering, processing, and analyzing questionnaires, as well as generating and distributing feedback reports. The content of the questionnaire items, feedback reports, and workshop could be developed within the organization. This approach will result in a questionnaire and process that is very tailored for the organization, but the organization would need to dedicate professional staff time and resources for the development.

Constructing the Multisource Feedback Questionnaire

How to Identify Competencies for a Multisource Feedback Process?

Multisource feedback questionnaire items should measure clearly defined behavioral competencies that are important for successfully performing the manager's job. Exhibit 8.1 provides an example of such competencies that have been used as a basis for a multisource feedback process.

One very good source of competencies is a well-developed, organization-specific managerial competency model that is the result of a sound job analysis and content validation process. This type of model will provide behavioral definitions that can be used as the basis for developing behaviorally based multisource feedback questionnaire items relevant for the current and future direction of the organization. The competency model should contain a core set of general managerial competencies, as well as separate sets of specific competencies for disciplines such as sales, finance, marketing, engineering, and so forth. An advantage of using a competency model is that its purpose is to form the foundation of the organization's HR systems, such as selection, performance management, development, promotion, compensation, and succession planning. As Sparrow and Bognanno (1993) point out, linking the multisource feedback process to it will horizontally integrate this process with other HR systems, creating consistency, coherence, and mutual reinforcement across the organization's HR practices. Sparrow and Bognanno note that linkage to the competency model will also result in vertical integration, tying this HR practice to business strategy by

Exhibit 8.1. Example of Competencies That Have Been Used as a Basis for a Multisource Feedback Process.

Competencies

- *Communication.* Clearly articulating performance expectations to staff, providing them with all relevant information to do their jobs, and creating an environment where candid communication is the norm.
- *Motivation.* Motivating group members by rallying them around a common goal, highlighting episodes of good performance, being considerate of their personal needs, and creating an environment where collaboration is the norm and where each group member feels valued.
- *Participation and Empowerment.* Encouraging and valuing the input and feedback of subordinates, providing the opportunities and resources for them to make decisions and accomplish work on their own, and supporting them in their decisions.
- *Performance Management.* Providing meaningful and timely performance feedback to direct reports, encouraging them to take advantage of personal development opportunities, and utilizing their input when setting each individual's objectives.
- *Planning.* Anticipating the resources needed to meet group or team objectives and ensuring understanding of short-term goals and long-term strategic direction.
- *Problem Solving.* Being accessible and willing to share knowledge and expertise to solve problems and resolve conflicts and disagreements when necessary.
- *Quality.* Stating and demonstrating commitment to quality and customer satisfaction and encouraging others to do so.
- *Valuing Diversity.* Demonstrating the willingness and ability to effectively manage groups consisting of diverse individuals and capitalizing on that diversity to achieve group goals.

Source: Copyright © 1992, Assessment Alternatives, Inc. Adapted from Assessment Alternatives, Inc., *Survey of Leader Behavior;* used with permission.

assessing important competencies required for the success of the business. Besides using organization-specific competency models as a basis for developing multisource feedback items, generic management competencies developed by outside consultants can be used, taking into account the considerations described earlier.

How Should the Multisource Questionnaire Be Constructed?

The multisource feedback items should describe observable behaviors, so that participants can use the resulting feedback from raters to behave more effectively. Yukl and Lepsinger (1995) suggest that the items should describe positive rather than negative behaviors, because feedback on negative behaviors may make managers feel more defensive, and does not tell participants what they should be doing, only what not to do.

The various sources may need to have somewhat different versions of the questionnaire items because they have the opportunity to view different behaviors. This may be especially true for rating sources outside the organization. However, for ease of administration, one version of the questionnaire is typically developed, with a rating option of "does not apply." London and Smither's (1995) survey results are consistent with this conclusion, with only three of the twenty respondents indicating that they sometimes develop customized versions for different sources.

Length of the questionnaire is an important consideration because it can affect response rate, especially if raters are completing questionnaires for more than one participant (Bracken, 1994). Experience suggests that this situation becomes especially important when all the managers in an organization of one hundred or more are participating, and direct report, supervisor, and peer ratings are being collected. For example, some managers may be asked to complete as many as fifteen or more questionnaires in such applications. Under these circumstances, a central coordinator may need to monitor and limit the number of questionnaires sent to any single rater. To facilitate responding, Bracken (1994) suggests that instruments contain forty to sixty items. However, where possible, even shorter questionnaires may be required to facilitate responding.

In the questionnaire instructions, raters should be assured of the confidentiality of their responses by stating the minimum number of respondents needed for data to be reported separately for the source. In their survey, London and Smither (1995) report that the typical number of raters required to present feedback for a source is three to five. Timmreck and Bracken's (1995) data are somewhat more skewed toward a minimum of three raters, with 57 percent of the processes requiring a minimum of three raters, 21 percent a minimum of four raters, and 7 percent a minimum of five raters. If the number of respondents falls below the minimum, then item data for this source are combined with another source that results in meeting the requirements for the minimum number of raters. Supervisors are the only exception. Their data are reported individually (if the participant lists only one supervisor) as would be the case with a performance appraisal.

What Types of Response Scales Should Be Used?

Research cited by Bracken (1996, 1994) recommends the use of "agreement scales" with anchors that range from "strongly agree" to "strongly disagree," or "satisfaction scales" with responses ranging from "very satisfied" to "very dissatisfied." Timmreck and Bracken's (1995) survey data indicated that "extent scales" with anchors ranging from "a great extent" to "a little extent" were most often used followed by the "agreement" and "satisfaction" scales. Research discussed by Bracken (1996, 1994) has also suggested that "frequency scales" with anchors such as "always," "usually," "sometimes," and "never" have some serious deficiencies. Much more systematic research is needed, however, to determine which types of response scales produce the most psychometrically sound multisource feedback data. Based on current research and practice, either agreement scales, satisfaction scales, or extent scales would be reasonable to use.

Some multisource instruments include, or have the option to include, an "expected frequency of performance scale" (anchors range from "never" to "always") and an "importance scale" (anchors range from "not important" to "very important"). Survey data from Timmreck and Bracken (1995) indicate that these scales are not often used. Including such scales is generally not recommended, because the items on the questionnaire should all be relevant for effective job performance, and so should all be rated relatively high on these scales (Dalessio, 1996). However, "importance scales" may be useful for identifying job-relevant items to be included in feedback reports when generic or off-the-shelf questionnaires are being used.

Consideration should also be given to the number of points on the rating scale, so that enough variance is produced to detect meaningful changes in behavior over time (Bracken, 1994), and meaningful differences between rating sources. Timmreck and Bracken (1995) report that rating scales ranging from four to seven points were most often used, and appear to afford reasonable choices.

Should Verbatim Comments Be Gathered?

Verbatim comments (or write-in comments) can provide the participant with very useful additional information that will be helpful with the interpretation of ratings, and are recommended as part of most multisource feedback processes. These comments are often collected as indicated by Timmreck and Bracken's (1995) survey, where 66 percent of the processes included verbatims. Verbatim comments can be collected by asking the rater at the end of the questionnaire to summarize the participant's key strengths and provide comments on how the participant could be more effective. Another approach to verbatim comments is to furnish the rater with the opportunity to write comments on ratings for each item. Although some raters may like this option, others may feel it is too cumbersome, and it could result in the participant receiving very few, or poor quality, verbatim comments (Bracken, 1996, 1994; Dalessio, 1996).

Verbatim comments are typically typed and included as a section in the feedback report, or attached to the report. Usually deleted from the comments are the names of other employees, as well as any expletives that may appear. The confidentiality of raters providing verbatim comments needs to be protected. Confidentiality becomes a real issue when relatively few raters respond to the multisource questionnaire. Under these circumstances, the participant may know (or believe he or she knows) who provided the comments. An approach that can be used in this situation is to content-analyze the comments and produce generic statements to the same effect. The relevant generic comments can then be fed back to the participant.

One potential downside to the use of verbatim comments is that some participants may focus on one or two comments that are not representative of the larger body of quantitative feedback. This misplaced focus can result in dwelling on a negative comment and ignoring the preponderance of positive ratings from raters who did not offer any written comments. This potential problem can be handled in two ways. First, by asking raters to provide comments on strengths as well as developmental needs, and second by having a skilled facilitator discuss this issue as part of the feedback workshop and warn participants not to focus too much on any single verbatim comment, but rather weigh all the information in the feedback report.

Administration of the Multisource Feedback Process

How Should an Organization Choose Raters?

A number of issues must be considered when choosing raters for a multisource feedback process. Research discussed earlier provides several suggestions for current best practice that include maintaining the anonymity of the raters (Antonioni, 1994) and obtaining feedback from as many qualified sources as possible (Bernardin et al., 1993; Conway & Huffcutt, 1996). As discussed earlier, research is not available on how the quality and acceptance of ratings is affected by participants choosing their raters versus others providing input. The suggestion based on current practice is participants should play the major role in selecting raters to promote empowerment and acceptance of the results. Supervisor input or review of the selected raters becomes important if the results are to be shared with the supervisor.

Another relevant issue is the number of raters that participants should be allowed to select. Allowing selection of an unlimited (or very large number) of raters can result in giving questionnaires to raters who have a limited opportunity to view the participant's behavior, and can affect the quality of ratings (Bracken, 1994). Selection of all direct reports typically is important so that the participant receives a complete picture from this source. However, this practice may not be possible in some line organizations where supervisors can have as many as thirty to forty direct reports. In these situations, a smaller number of direct reports (for example, ten to fifteen) may need to be randomly selected for participation. Random selection allows for the opportunity to receive differing views, yet not have others feel undervalued. All direct reports should be informed that random selection has been used in this situation. If the participant has as many as ten to twelve direct reports, all should receive the questionnaire, then the combined number of peers and internal customers sent the questionnaire may need to be limited to a total of approximately six to eight for practical reasons. When there are very few or no direct reports, the combined number of peers and internal customers sent the questionnaire could be increased to approximately ten to fifteen. Numbers at the higher end of these ranges should be selected if lower response rates are expected.

If external customers or vendors are included, similar limits may also need to be required. Including these sources complicates the process, because different questionnaires may need to be structured for them, and motivation to provide high-quality input may not be the same as for employees (Bracken, 1994). The participant and the supervisor should also complete a questionnaire. In their survey, London and Smither (1995) indicated that the typical number of raters from which ratings were obtained ranged from three to nine. Depending on the response rate, the ranges suggested here would probably result in a larger number of respondents.

Yukl and Lepsinger (1995) suggest that raters should be chosen who have interacted with the participant for a year or more on a regular basis. This is good advice, because if raters have not had enough opportunity to observe the participant, more ratings may show central tendency, and more items may not be rated.

What Type of Rater Training Should Be Provided?

Timmreck and Bracken's (1995) survey indicated that 66 percent of the multisource feedback programs included rater training. Bracken (1996, 1994) and Yukl and Lepsinger (1995) have suggested that rater training is a useful component for improving the quality of a multisource feedback process, and should be provided in some form. Such training could include information on the purpose of the process, instrument design, questionnaire administration, as well as training on behavioral observation, rater errors, and the psychometric quality of the ratings. Smith (1986) concluded that rater training processes where raters become more actively involved in the training, for example, participating in group discussions along with practice and feedback exercises, produce the best results with regard to rater errors and psychometric quality. Chapter Eleven in this volume addresses these issues in more depth.

The type of rater training discussed in this section would require bringing raters together for facilitated meetings. The participants, who will also serve as raters for the process, could possibly complete rater training as part of the communication process. However, given that some potential raters may be in different units within the organization—or outside the organization—such as internal customers and external customers, training all raters may not be logistically possible. At a minimum, rater training information could be provided to these raters in the format of a short memo or booklet.

How to Distribute and Return Questionnaires?

Distribution and return of questionnaires is a major administrative task that may positively or negatively affect participants' attitudes toward the multisource feedback process depending on how it is handled. To be effective, raters need to feel that the distribution and return of questionnaires protects the anonymity of their responses, and participants must feel that they are receiving feedback from as many raters as possible in a timely fashion. Outside consultants provide a very useful service in the distribution and return of questionnaires, because they can act as a neutral party and maintain the anonymity of raters during the process.

Three main activities need to be accomplished in this process: participants need to identify their raters for the administration; raters need to receive questionnaires and participants need to receive their self-questionnaire; and raters and participants need to return their questionnaires by the due date for analysis. The main methods for administration and distribution are mailing (either paper-and-pencil questionnaires or questionnaires on 3.5-inch computer disks); electronic mail via local area networks, the company intranet, or the Internet; or an automated telephone system.

When the mail is used, the names and addresses of participants can be given to the vendor by a company administrator. The vendor can then send out a form through the company mail to participants, asking them to list the names, addresses, and phone numbers of their raters. The vendor can then assemble a packet for each selected rater that contains all the multisource questionnaires for the participants that they need to evaluate, along with the self-questionnaire (if appropriate). The packet can then be distributed through the company mail. This is an efficient way to distribute the questionnaires, because all raters and participants receive only one packet. Within the packet is a generic cover letter explaining the process and a postage-paid return envelope addressed to the vendor via U.S. mail. Such a process saves costs by completing the initial distribution through company mail and the more sensitive return through U.S. mail.

Another method that uses the mail for questionnaire distribution involves having the vendor mail each participant a packet that contains a self-questionnaire and a specified maximum number of questionnaires for the participant to distribute. Each questionnaire would need to be enclosed in an envelope that could be sent through company mail to the rater. A generic cover letter explaining the project and a U.S. postage-paid return envelope addressed to the vendor also would need to accompany each questionnaire in the envelope. A problem with this latter method is that raters do not receive a single packet of questionnaires, but rather individual questionnaires from each participant, and no central control can be put into place to limit the questionnaires given to a single rater.

With both mail methods, the vendor would need to know the number of raters selected by the participant from each source so that a tally of the number of responses can be kept. Prior to the return date, each participant should be notified by the vendor or a company administrator of the number of returns for each participant for each source. If less than a minimum number of raters has returned questionnaires for a source, the participant then needs to e-mail or voice mail all participants to ask them to return questionnaires as soon as possible.

A promising alternative to the mail process is distribution and return of questionnaires via e-mail through a local area network, the company intranet, or the Internet. This type of administration has the advantage of speeding the questionnaire distribution and return process, easing the modification of the questionnaire, and facilitating the data coding for analysis. Potential challenges for such a system are securing the transmission of data and maintaining and ensuring the anonymity of the raters. Also, such systems can be expensive to develop.

A final alternative distribution and return system is through use of an automated phone system. With such a system, participants and raters can enter their questionnaire ratings through following a menu activated by a call-in number. Participants can check the tally of the number of respondents and request the generation of a report when they wish. Like an e-mail system, the process can speed data collection, coding, and analysis, but can be expensive to develop.

Constructing Feedback Reports

How Should Data Be Aggregated for Presentation?

Participants in multisource feedback processes receive a feedback report summarizing ratings from the various sources on the items and competency dimensions, as well as verbatim comments. Practitioners will need to make decisions regarding options for aggregating data and for the content of various sections of the feedback report.

Mean ratings are typically calculated for reports. For example, in the Timmreck and Bracken survey (1995), 76 percent of the multisource processes reported presenting means on the feedback report. Average ratings for each source are often calculated on items and dimensions. Overall mean ratings that include data from all sources (except the self-ratings) are also commonly calculated. Practitioners responsible for designing the process will need to decide whether all individual raters should be equally weighted when calculating these averages, or whether the rating sources should be equally weighted. Equally weighting the sources will give more weight to supervisor ratings and sources with fewer raters.

Weighting all raters equally is the preferred method, so that the participant's interpretation of the overall rating is not overly influenced in particular by the supervisor's rating. The supervisor's rating usually is presented as a separate source in the feedback report, so the participant has an opportunity to understand the supervisor's perspective. Furthermore, participants receive input from supervisors at annual performance appraisal reviews. Multisource feedback is an opportunity to emphasize information from other sources.

Olympic scoring—where the highest and lowest ratings are eliminated before calculating means—is another option for consideration. Bracken (1996) has strongly suggested that Olympic scoring is not appropriate because it may eliminate useful data from raters with a large amount of information on the participant. Olympic scoring is inconsistent with the assumption that participants need to consider input from all raters to better understand their behavior.

Including an index of item variance within the source (for example, range or standard deviation) can help participants interpret their evaluations, and may be a better approach than Olympic scoring to addressing variation and degree of agreement in ratings. The extent to which indices of item variance are used on feedback reports is unclear. Only 17 percent of Timmreck and Bracken's (1995) multisource processes indicated that a range was included on the feedback report, while 70 percent of the London and Smither (1995) respondents indicated providing an indicator of within-source agreement. The range is probably the best method for summarizing within-source agreement, because it is easiest to understand in comparison to the standard deviation, for example. Strong consideration should be given to providing an item variance index on the feedback report. However, this consideration needs to be balanced against possibly overwhelming participants with so much information that they are not able to use the report effectively to identify their major strengths and areas for development. If an index of item variance is added to a report, perhaps another aspect of the report should be streamlined.

How Should Feedback Report Sections Be Presented to Facilitate Interpretation?

The purpose of any feedback report section is to assist the participants with interpreting their results and gaining insight into their behavior. The various report sections should provide the participants with views of their behavior from each of the different sources, and with feedback regarding strengths and developmental needs that when considered together will lead to a complete picture and quality conclusions. Most report sections present summaries of numerical ratings, but as discussed earlier a section of verbatim comments is often included.

One typical report section presents overall average ratings (excluding self-ratings) for the competency categories listed from the highest to lowest rated. This section provides a quick summary picture of how others view the participant's strengths and developmental needs on the broad competency categories, and can help identify common themes among competency categories that are viewed as strengths as well as developmental needs. Generally, these overall mean ratings, as well as mean ratings for the individual sources, tend to be skewed in such a way that they are above the scale midpoint for most participants. Participants need to be cautioned that the scale midpoint in most cases should not be used as a benchmark for average performance; a scale value above the midpoint is generally a better benchmark. A consensus of subject matter experts could be used to define the benchmark. Participants need to understand that they should use this information just as a benchmark and consider their overall pattern of results, trying to focus on their relative top strengths and developmental needs in action planning.

Another report section will list from five to ten of the highest and lowest overall average rated items. The competency dimension under which the item is categorized can be listed next to each item. This section can help participants identify more specific behaviors within each of the competency dimensions that are strengths and developmental needs, as well as allow participants to determine whether the highest or lowest rated items are concentrated in one or two dimensions, or are spread across dimensions.

After the report sections focusing on the presentation of overall average ratings (excluding self-ratings), other sections of the report typically compare item and competency dimension information for each of the rating sources. They present information from the self-rating and from supervisor, direct reports, peers, and customers, as well as an overall average. Results of the London and Smither (1995) survey indicated that 80 percent of their respondents contrasted ratings from different sources on the same page and 90 percent explicitly contrasted self- with other ratings. An example of a report section presenting these types of comparisons for self-, supervisor, direct reports, and overall average ratings is shown in Exhibit 8.2.

Graphic displays are usually presented at the competency dimension level and allow for quick visual comparisons of similarity and differences among the various sources. A graphic display may show, for example, key competencies where the sources agree or disagree in their ratings, or may show that the various sources perceive the same pattern of strengths and developmental needs, but

Exhibit 8.2. Example Report Comparing Self-, Supervisor, Direct Report, and Overall Ratings on Items and Dimensions.

Behavioral Statement Ratings by Performance Category

In this section, you can compare your Self-Assessment for each individual statement about leadership behavior that appeared in the survey to the ratings made by your supervisor, your direct reports, and the overall average of ratings.

(1 = Not at all, 3 = To an average extent, 5 = To a very great extent)
Your "Overall" rating is based on 11 people.

	Self	Supervisor	Direct Reports	Overall
Communication	**3.7**	**3.3**	**2.8**	**2.9**
__ Clearly stated expectations regarding team's performance.	3.0	3.0	2.6	2.7
__ Provided direct reports with all relevant information necessary to effectively do their jobs.	4.0	3.0	2.4	2.5
__ Encouraged open and honest communication among team members.	4.0	4.0	3.5	3.7
Motivation	**3.8**	**3.0**	**3.0**	**3.0**
__ Encouraged and facilitated teamwork and collaboration.	3.0	4.0	2.9	3.2
__ Rallied team members around a common goal.	4.0	3.0	2.8	2.8
__ Made all team members feel valued.	4.0	3.0	2.6	2.6
__ Let direct reports know when they did a good job.	4.0	2.0	3.3	3.3
__ Were considerate of the personal needs of direct reports outside the workplace.	4.0	3.0	3.2	3.1
Participation/Empowerment	**3.8**	**3.0**	**3.4**	**3.3**
__ Allowed direct reports to participate in making decisions that affected them.	4.0	3.0	3.0	3.0
__ Empowered direct reports to make decisions on their own.	4.0	4.0	3.4	3.5
__ Assisted direct reports in obtaining the necessary resources to effectively do their jobs.	4.0	3.0	3.9	3.7
__ Supported the commitments direct reports made to their customers.	4.0	2.0	2.9	2.8
__ Requested and acted on feedback from direct reports about ways to improve your effectiveness.	3.0	3.0	3.4	3.3
Performance Management	**3.5**	**2.5**	**2.9**	**2.7**
__ Provided direct reports with coaching and guidance on ways to improve their effectiveness.	3.0	3.0	3.1	3.0
__ Provided direct reports with timely, specific feedback on their performance.	3.0	3.0	3.5	3.4
__ Informed and encouraged my direct reports to take advantage of personal development opportunities.	4.0	2.0	2.6	2.4
__ Encouraged and utilized each direct report's input when setting his or her objectives.	4.0	2.0	2.5	2.3
Planning	**4.0**	**3.3**	**2.4**	**2.6**
__ Clearly communicated short-term goals and long-term strategic direction.	4.0	4.0	2.5	2.7
__ Planned in advance and anticipated resources that would be needed to meet the team's objectives.	4.0	3.0	2.4	2.6
__ Clearly explained to each direct report how his or her individual function relates to long-term strategic direction.	4.0	3.0	2.4	2.5
Problem Solving	**4.3**	**3.3**	**4.0**	**3.7**
__ Willingly shared knowledge and expertise with all direct reports.	5.0	3.0	4.5	3.8
__ Were readily accessible if direct reports needed to discuss a problem or a particular issue.	3.0	3.0	3.8	3.5
__ Took steps to resolve conflict and disagreement within the team.	5.0	4.0	3.6	3.8
Quality	**4.0**	**4.0**	**4.1**	**4.1**
__ Regularly challenged direct reports to continuously improve their effectiveness.	4.0	4.0	3.7	3.6
__ Encouraged direct reports to use creative and innovative ways to satisfy the customer.	4.0	4.0	3.8	4.1
__ Demonstrated a strong commitment to customer satisfaction with day-to-day actions.	4.0	4.0	4.3	4.1
__ Regularly stressed the importance of quality and continuous improvement.	4.0	4.0	4.3	4.3
Valuing Diversity	**3.8**	**3.5**	**3.7**	**3.7**
__ Encouraged and accepted points of view that differed from your own.	4.0	3.0	3.6	3.6
__ Utilized the expertise of all direct reports when appropriate.	4.0	4.0	3.8	3.9
__ Valued the contributions of all individuals, regardless of level.	4.0	4.0	3.6	3.7
__ Utilized the diverse knowledge, skills, and abilities of team members.	3.0	3.0	3.8	3.6

Source: Copyright © 1992, Assessment Alternatives, Inc. Adapted from Assessment Alternatives, Inc., *Survey of Leader Behavior;* used with permission.

at different levels of effectiveness. Numerical displays are used effectively to compare ratings from the various sources at the item level, which allows the participant to see how the sources differ in terms of ratings of specific behaviors.

When ratings from these displays show that all sources are high and agree, this is a strength to be leveraged. When all sources agree, and the rating is low, this is a priority for developmental efforts. When some sources rate high and agree, while other sources agree and rate low, this may be a behavior that is displayed in some situations and not others. The participant may need to consider how to display this behavior more consistently. If all sources agree and rate low, but the self-rating is high, the participant should keep in mind that self-ratings tend to be lenient (for example, Harris & Schaubroeck, 1988) and this may be an area for developmental planning.

Norms should be presented for items and competency dimensions when available. Norms show how the participant compares to other managers, and will further facilitate interpretation (Yukl & Lepsinger, 1995). Norms can be based on national samples, or large samples from the organization. National norms will be available from vendors that administer multisource feedback in a wide variety of organizations. The quality of these norms should be evaluated in terms of sample size, relevance of managerial level, and relevance of industry. Norms from the organization are also useful, but an adequate sample size on which to base these norms may take a long time to gather.

A final report section could provide suggestions for on-the-job developmental activities, readings, and seminars relevant to the key behaviors for each of the lowest-rated competency dimensions. Rather than actually being placed in the report, this information could be consolidated into a separate booklet with all the competency dimensions included. The participant could use this booklet as a reference for obtaining developmental suggestions for any dimension. An example of a page providing on-the-job suggestions for two competency areas is shown in Exhibit 8.3.

How Should the Department (or Group) Feedback Report Be Used?

Often a department-level or group-level feedback report will be produced as part of the multisource feedback process to help identify group strengths and developmental needs for the purpose of

Exhibit 8.3. Developmental Suggestions for Two Competencies.

Suggestions for Your Development

This section provides several specific recommendations for improving your performance in each performance category that has been listed as a "developmental opportunity." Remember, these are only suggestions to guide your own developmental efforts.

It is important to keep in mind that an indication of developmental opportunity in a particular category does not imply that you have *no* skill in that area. By the same token, an indication of strength in a performance category does not imply that you should not focus on maintaining or improving your performance in that area. The purpose of this report is to guide your efforts as you continuously monitor and improve your performance in *all* areas.

Performance Management

- Start keeping a performance diary by writing down positive or negative examples of performance by your subordinates. Share the feedback with subordinates in a constructive, nonthreatening way at regular intervals.

  Set a goal to review each subordinate's performance once every two weeks and provide developmental feedback the next day.
- Keep an updated list of developmental opportunities for your subordinates and actively encourage them to take advantage of the activities at regular intervals.

  Encourage subordinates to actively seek feedback from you.
- Develop a plan for assessing employee needs for coaching. Have each employee prepare a list of areas in which he or she feels coaching is needed. Meet individually with each employee and develop an agreement as to what coaching will be provided.
- Plan to meet with each subordinate at regular intervals for the purpose of giving feedback. Make feedback clear and specific and link feedback to objectives.
- When delivering negative feedback to your subordinates, focus on maintaining their self-esteem at all times. Keep the discussion focused on the behavior you have observed rather than your overall impressions.
- Allow employee input in setting objectives. Work toward mutually agreed-upon goals for the employee and review progress at regular intervals.

Exhibit 8.3. Developmental Suggestions for Two Competencies, cont'd.

Planning

- Solicit employee input on resource problems. Work with employees to develop methods to ensure that resources are not a problem.
- With employees, develop a plan to ensure that any commitments made to customers are realistic and are communicated clearly to all relevant parties. Ensure that the resources necessary to meet those commitments are available.
- Formally communicate short-term goals and long-term strategic direction via written documents or video presentation.
- Check with each subordinate to see that he or she understands short-term goals and long-term strategic direction.
- Take personal interest in developing a career path for each of your subordinates. Regularly monitor progress and provide feedback and suggestions for aligning current responsibilities with long-term goals.
- Explain to each subordinate, informally in conversation and formally in goal-setting, how his or her specific function relates to long-term strategic direction.

Source: Copyright © 1992, Assessment Alternatives, Inc. Adapted from Assessment Alternatives, Inc., *Survey of Leader Behavior;* used with permission.

departmental goal setting. The sections of this report will parallel those of the participant report, but will be based on average ratings for all participants in the group. For example, this type of report could be used to identify whether employees in a group of succession candidates are managing well upward with their supervisors, but less well with peers, customers, and direct reports. If such trends are apparent, then the group head can develop appropriate action plans together with the participants. The group report can be shared with participants during the feedback workshop and used as a stimulus to practice goal setting and action planning. Suggestions resulting from the discussion could be considered for implementation.

Facilitating Support for Developmental Activities

What Should Be Included in a Facilitated Feedback Workshop?

The purpose of the feedback workshop is to provide participants with an opportunity to learn how to interpret their feedback reports, and begin to write follow-up action plans to improve in areas that are opportunities for development and leverage areas that are strengths. Workshop facilitators can be more effective in achieving these purposes if they have completed a multisource feedback process themselves, so that they can speak about the various aspects of the process from direct experience.

The feedback workshop can begin with a discussion of the purpose, objectives, value, and benefits of multisource feedback. Next, background should be provided about how the multisource feedback instrument was developed and the competency model on which it is based. The workshop provides an opportunity to communicate and clarify critical organizational competencies. Then protection of the anonymity of the raters should be discussed, along with the confidentiality of the results. Who does (and does not) have access to the results needs to be clarified, and reasons provided for access.

The information in the various sections of the feedback report should be explained next, so that participants understand their data. Also, various interpretations of discrepancies among the rating sources should be discussed so that participants can use their data effectively for action planning. As discussed earlier, research by Bass and Yammarino (1991) and Van Velsor et al. (1993) suggests that as part of the feedback workshop facilitators should focus the attention of participants on large self-other discrepancies, especially when self-ratings are consistently high. These participants should consider working to bring their self-views more in line with those of others. Research by Fleenor et al. (1996) emphasizes the need for facilitators and participants to consider not only discrepancies but also behaviors receiving low to average ratings as opportunities for development. Yammarino and Atwater (1997) present a model that takes into account both the degree of discrepancy between self-other ratings and the level of the rating. The model presents four different self-other agreement categories, each of which is associated with different possible consequences. The categories are *overestimators* (who tend to rationalize negative feedback and may be headed for career derailment), *in-agreement/good* (who are effective managers and leaders), *in-agreement/poor* (who have accurate self-perceptions, but are poor performers and currently may be taking little action to improve performance), and *underestimators* (who do not appear to recognize their strengths and may not be realizing their full potential). Workshop facilitators can use the model to illustrate why different patterns of results require action to be taken by participants. Facilitators need to emphasize that self-other discrepancies point out behaviors that are salient to others, and need greater awareness and consideration on the part of the participant in developmental planning. Acceptance of such feedback can often be difficult for participants.

The work of Ashford and Tsui (1991) also indicates that facilitators should suggest that participants be more open to seeking negative feedback about their behavior from supervisors, direct reports, and peers as part of the developmental process. Finally, participants should be directed to use the information in the report for identifying strengths as well as developmental needs. Participants more naturally focus on developmental need areas, but they should be reminded to also consider their strengths so that these can be leveraged where possible to further increase effectiveness.

In the workshop, participants should also be provided with guidance on how to develop written action plans that direct their effort toward a specific goal. Facilitators need to discuss topics such as prioritizing strengths and developmental needs, setting specific goals, analyzing and choosing developmental activities, developing an action plan with time frames, taking action, and measuring results. Goals should be set for the highest-priority developmental needs. The goals that are set in these areas should be specific and challenging yet attainable. The goals also should be consistent with the mission of the organization. Time should be set aside during the workshop to actually begin setting goals and writing developmental plans. The participants should be left with a workbook that covers the topics in the workshop for reference purposes (Bracken, 1994).

What Follow-Up Activities Should Be Available?

Once a multisource feedback process is completed, participants need to take the information from the feedback workshop and incorporate it into follow-up developmental activities. Hazucha et al.'s (1993) research indicated that self-ratings and others' ratings of effort to change following the implementation of multisource feedback were related to whether the organization was reported as having a formal and active career development program. Organizations may vary to a large extent in terms of whether multisource feedback processes are linked to career development programs. Results from London and Smither's (1995) survey showed that only 40 percent of their respondents reported that multisource feedback is linked to a specific development program. Linking multisource feedback to these programs may be one key to gaining the greatest benefit from the process.

A process put into place by a company in the financial and insurance industry provides a good example of how multisource feedback can be linked to a development program. In this organization, managers were rated by their direct reports and peers on a few key organizational competencies prior to attending a weeklong development program. On each morning of the program, the participants received the multisource feedback relevant to the competency that would be developed during the remainder of the day. For example, if the day's training focused on conflict management and negotiation skills, then the participants would start the day by receiving feedback about these competencies first thing in the morning. At the end of day, the participants set goals and action plans in these areas.

On a similar issue, Yukl and Lepsinger (1995) suggest that the benefits of multisource feedback are more likely to be achieved if other follow-up activities also support the process. These activities include not only supporting formal training sessions, but also on-the-job learning opportunities for competencies identified as developmental needs. Also, participants should be provided with the opportunity to obtain on-the-job coaching from supervisors or peers on identified strengths and developmental needs. Follow-up administration of the multisource feedback questionnaire can also provide participants with information on what changes in behavior are being perceived. The London and Smither (1995) survey suggested that about 60 percent of their organizations administered the process to participants more than once. Finally, the Walker (1997) study suggests that holding feedback meetings with subordinates may be an important follow-up activity to assist improving performance.

Evaluation and Revision of the Process

Program evaluation should be built into the process to determine whether the multisource feedback is achieving its objectives. For example, the process may include objectives such as: having all employees design a formal written developmental plan, incorporating specific achievable developmental activities from the plan into performance appraisal objectives, or increasing involvement of employees in formal development programs. Data could be collected to evaluate the extent to which these objectives had been achieved.

As part of the evaluation, employee reactions to the process could also be solicited through a questionnaire or focus groups. Information should be sought to improve major aspects of the process, including adequacy of communications, definitions of competencies, questionnaire content, administration process, usefulness of feedback reports, feedback workshops, and adequacy of developmental activities. Finally, longitudinal studies could also be conducted to evaluate the extent to which improvement in performance is affected by the multisource feedback process. As mentioned earlier, these studies are often difficult to conduct because of complex methodological issues. Based on the results of the program evaluation, changes could be made that include administering it more frequently, stopping it, requiring more accountability in writing developmental plans, or providing more support for developmental activities.

Summary and Conclusions

Exhibit 8.4 provides a summary of recommendations for implementing multisource feedback processes based on current research and practice.

Multisource feedback processes are popular means for providing managers with helpful developmental feedback. The extent to which the process is successful in an organization depends to a large extent on how well it fits with the strategy and culture of the organization and is supported by a set of mutually reinforcing HR practices. As discussed previously, one approach to accomplishing

Exhibit 8.4. Summary and Recommendations for Implementing Current Multisource Feedback Practices.

Planning the Multisource Feedback Process

Linking the Process to Organizational Objects

- Base the process on important organizational competencies identified as key to the organization's success.
- Link the process to critical organizational change efforts.

Gaining and Displaying Senior Management Support for the Process

- Plan data collection to demonstrate improvements in performance as a result of the process.
- Define outcomes of the process in measurable terms, and summarize measures.
- Cascade the implementation of the process from the senior level down.
- Have senior managers feedback their results to their direct reports to act as role models and supporters of the process.

Using Multisource Feedback for Development or Personnel Decisions

- Use validated competencies as the basis for a developmental or decision tool.
- Address the issue of holding participants in a purely developmental multisource feedback system accountable for completing development planning and activities.
- Address the issues of quality of ratings and participant reactions in a multisource feedback system used as input to personnel decisions.
- Note that multisource feedback processes are generally suggested for use as developmental tools.

Providing Up-Front Communications

- State clear reasons why the process is being used for development or personnel decisions.
- State how rater anonymity will be maintained.
- Provide information on plans for scheduling, questionnaire content, frequency of feedback, follow-up support, and how long the process will be in place.

Working with Outside Consultants

- Request psychometric data on vendor's items for the purposes of evaluating, selecting, and modifying items.
- Request and evaluate vendor's norms for purposes of feedback.

Exhibit 8.4. Summary and Recommendations for Implementing Current Multisource Feedback Practices, cont'd.

- Use outside consultants to distribute, return, and analyze questionnaire data so that rater anonymity is ensured.

Constructing the Multisource Feedback Questionnaire

Identifying Competencies

- Base the questionnaire on an organization-specific competency model that is linked to other HR systems and organizational strategies.
- Establish job-relatedness through a content validation process for programs that use generic vendor-developed competencies.

Constructing the Questionnaire

- Have each item describe observable behavior linked to job-relevant competencies.
- Limit questionnaire length to forty to sixty items if at all possible.

Choosing a Response Scale

- Use "agreement," "satisfaction," or "extent" scales with four to seven points, bearing in mind that more research on the psychometric soundness of scales is needed.
- Avoid "expected frequency of performance" scales, as they will not provide much information if all questionnaire items are job-relevant.

Verbatim Comments

- Include verbatim comments as a supplement to ratings in most multisource feedback processes.
- Collect verbatim comments by asking raters at the end of the questionnaire to summarize strengths and areas for development.

Administration of the Multisource Feedback Process

Choosing Raters

- Have participants choose raters with whom they have interacted for a year or more on a regular basis.
- Maintain rater anonymity.
- Have participants complete a self-questionnaire.
- Send questionnaires to ten or twelve direct reports. With larger numbers of direct reports, random selection should be used—and communicated.
- Choose ten to fifteen peers or customers when the participant does not have direct reports. When there are direct reports, peers or internal customers may need to be limited to six to eight.

Exhibit 8.4. Summary and Recommendations for Implementing Current Multisource Feedback Practices, cont'd.

Providing Rater Training

- Provide rater training in group settings where possible, or through written materials.

Distributing and Returning Questionnaires

- Accomplish these activities: participants identify raters; raters receive questionnaires; raters return questionnaires for analysis.
- Select one of these methods for distribution: mailing paper-and-pencil questionnaires or questionnaires on 3.5-inch disks; electronic distribution via local area networks, the company intranet, or the Internet; data collection through the automated telephone system.

Feedback Reports

Aggregating Data for Presentation in Report

- Weight all raters equally when calculating overall average ratings that include data from all sources (except self-ratings).
- Avoid use of Olympic scoring.
- Use the range if an index of within-source agreement is included.

Presentation of Feedback Report Sections to Facilitate Interpretation

- Present overall average ratings (excluding self-ratings) for the competency categories listed from highest to lowest.
- Present from five to ten of the highest and lowest overall average rated items.
- Present average item and competency ratings for each source.
- Use graphic display to present visual comparisons among various sources at the competency level.
- Present norms for items and competency dimensions where available.
- Provide suggestions for relevant on-the-job developmental activities, readings, and seminars in the report or as a separate booklet.
- Caution participants not to rely on the midpoint as a benchmark for average performance because ratings are typically skewed, but rather to use a scale value above the midpoint.

Using Department or Group Feedback Reports

- Present the group report in a form parallel to the participant report, but based on average ratings from all participants.

Exhibit 8.4. Summary and Recommendations for Implementing Current Multisource Feedback Practices, cont'd.

- Use the group report as a stimulus during the feedback workshop to practice goal setting and action planning.
- Use the group report to identify and implement developmental activities for the entire department.

Facilitating Support for Developmental Activities

Facilitated Feedback Workshop

- Use workshop facilitators who have completed a multisource feedback process themselves, so that they can speak from direct experience.
- Discuss purpose, objectives, value, and benefits of the feedback process.
- Provide background on competencies and questionnaire development.
- Discuss how rater anonymity will be maintained.
- Explain the sections of the feedback report, and how to interpret potential results.
- Provide participants with guidance on developing action plans and directing efforts to a specific goal.
- Leave participants with a workbook that covers workshop topics.

Follow-Up Activities

- Link multisource feedback to specific developmental programs.
- Support the multisource feedback process with on-the-job learning opportunities and coaching from supervisors or peers on strengths and developmental needs.
- Provide opportunities for follow-up administration of the process.
- Encourage participants to share their feedback results with subordinates in a group meeting.

Evaluating and Revising the Process

- Determine whether the process is achieving specific objectives such as designing a written developmental plan, incorporating identified developmental activities into performance appraisal objectives, or increasing involvement of employees in formal developmental programs.
- Assess employees' reactions and use this information to revise the program.
- Conduct longitudinal studies to evaluate whether the process affects improved performance.

this integration is through linking the multisource feedback process to a competency model that has been strategically designed to build skill sets required to successfully position the organization in new and changing markets, and is the foundation of other HR systems in the organization. Also, the multisource feedback process should be linked to specific company development programs and follow-up activities to facilitate and sustain the development of the participants. The value of the process also needs to be demonstrated by improved performance.

Only time will tell whether this methodology is here to stay, or will metamorphose into an alternative approach to feedback. Continued focus on validation issues, improved design of the multisource feedback programs, and research on the topic will result in a better process to improve self-understanding, human interactions, and effectiveness.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement Education. (1985). *Standards for educational and psychological tests* (3rd ed.). Washington, DC: American Psychological Association.

Antonioni, D. (1994). The effects of feedback accountability on upward appraisal ratings. *Personnel Psychology, 47,* 349–356.

Ashford, S. J., & Tsui, A. S. (1991). Self-regulation for managerial effectiveness: The role of active feedback seeking. *Academy of Management Journal, 34,* 251–280.

Atwater, L., Roush, P., & Fischthal, A. (1995). The influence of upward feedback on self- and follower ratings of leadership. *Personnel Psychology, 48,* 35–59.

Bass, B. M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology, 40,* 437–454.

Bernardin, H. J., Dahmus, S. A., & Redmon, G. (1993). Attitudes of first-line supervisors toward subordinate appraisals. *Human Resource Management, 32,* 315–324.

Bernardin, H. J., Hagan, C., Ross, S., & Kane, J. S. (1995, May). The effects of a 360-degree appraisal system on managerial performance. In W. W. Tornow (Chair), *Upward feedback: The ups and downs of it.* Symposium conducted at the Tenth Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Bracken, D. W. (1994). Straight talk about multirater feedback. *Training and Development, 48,* 44–51.

, D. W. (1996). Multisource (360-degree) feedback: Surveys for
dividual and organizational development. In A. I. Kraut (Ed.), *Or-
nizational surveys: Tools for assessment and change* (pp. 117–143). San
ancisco: Jossey-Bass.
S., Fleenor, J. W., & London, M. (1998). Elements of effective 360-
gree feedback. In W. W. Tornow & M. London (Eds.), *Maximiz-
g the value of 360-degree feedback: A process for successful individual and
ganizational development.* San Francisco: Jossey-Bass.
S., Fleenor, J. W., & McCauley, C. D. (1996, April). *Self-other rating
screpancy in 360-degree feedback: An investigation of demographic and
rsonality predictors.* Paper presented at the Eleventh Annual Con-
rence of the Society for Industrial and Organizational Psychology,
n Diego, CA.
Hampton. (1994). *Multi-rater assessments: Study of features and costs.*
ampton, VA: Department of Human Resources.
, J. M., & Huffcutt, A. I. (1996, April). *Testing assumptions of 360-
gree feedback: A meta-analysis of supervisor, peer, subordinate, and self-
tings.* Paper presented at the Eleventh Annual Conference of the
ciety for Industrial and Organizational Psychology, San Diego, CA.
, A. T. (1996). Review of *20/20 Insight (Version 2.1 for Windows).*
rsonnel Psychology, 49,* 1050–1056.
M. (1996). Multirater feedback and conditions for change. *Con-
lting Psychology Journal, 48,* 12–16.
eau, F., Jr., Graen, G., & Haga, W. J. (1975). A vertical dyad link-
e approach to leadership within formal organizations. *Organiza-
nal Behavior and Human Performance, 13,* 46–78.
ck, P. C., Reilly, R. R., & McGourty, J. W. (1997). The effects of
eer feedback on team member behavior. *Group and Organization
anagement, 22,* 508–520.
te, M. D. (1993). My hammer or your hammer? *Human Resource
anagement, 32,* 373–384.
L., Cannella, A. A., Jr., & Bedeian, A. G. (1991). Peer ratings: The
pact of purpose on rating quality and user acceptance. *Group and
rganization Studies, 16,* 367–386.
r, J. W., McCauley, C. D., & Brutus, S. (1996). Self-other rating
greement and leader effectiveness. *Leadership Quarterly, 7,* 487–506.
in, I. L., Zedeck, S., & Schneider, B. (1993). An exploration of the
b analysis–content validity process. In N. Schmitt & W. C. Borman
Eds.), *Personnel selection in organizations* (pp. 3–34). San Francisco:
ossey-Bass.
R. M. (1965). *Personnel testing.* New York: McGraw-Hill.
R. M. (1980). On trinitarian doctrines of validity. *Professional Psy-
ology, 11,* 385–398.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-s
self-peer, and peer-supervisor ratings. *Personnel Psychology,*
Hazucha, J. F., Hezlett, S. A., & Schneider, R. J. (1993). The impa
degree feedback on management skills development. *H
source Management, 32,* 325–351.
Kane, J. S., & Lawler, E. E., III. (1978). Methods of peer assessm
chological Bulletin, 85, 555–586.
London, M., & Beatty, R. W. (1993). 360-degree feedback as a
tive advantage. *Human Resource Management, 32,* 353–372
London, M., & Smither, J. W. (1995). Can multisource feedba
perceptions of goal accomplishment, self-evaluations, ar
mance-related outcomes? Theory-based applications and
for research. *Personnel Psychology, 48,* 803–839.
London, M., & Wohlers, A. J. (1991). Agreement between subor
self-ratings in upward feedback. *Personnel Psychology, 44,* 3
Mayfield, E. C. (1970). Management selection: Buddy nomin
visited. *Personnel Psychology, 23,* 377–391.
Mayfield, E. C. (1972). Value of peer nominations in predicti
surance sales performance. *Journal of Applied Psychology, 56*
Murphy, K. R., & Cleveland, J. N. (1995). *Understanding perfo
praisal: Social, organizational, and goal-based perspectives.*
Oaks, CA: Sage.
Nilsen, D., & Campbell, D. P. (1993). Self-observer rating disc
Once an overrater, always an overrater? *Human Resource M
32,* 265–281.
Reilly, R. R., Smither, J. W., & Vasilopoulos, N. L. (1996). A lo
study of upward feedback. *Personnel Psychology, 49,* 599–6
Romano, C. (1994). Conquering the fear of feedback. *HR Fo*
9–19.
Smith, D. E. (1986). Training programs for performance appr
view. *Academy of Management Review, 11,* 22–40.
Smither, J. W., London, M., Vasilopoulos, N. L., Reilly, R. R
R. E., & Salvemini, N. (1995). An examination of the ef
upward feedback program over time. *Personnel Psychology*
Sparrow, P. R., & Bognanno, M. (1993). Competency require
casting: Issues for international selection and assessmen
tional Journal of Selection and Assessment, 1, 50–58.
Society for Industrial and Organizational Psychology. (1987). *P
the validation and use of personnel selection procedures.* Lak
Graphic.
Timmreck, C. W., & Bracken D. W. (1995, May). Upward feed
trenches: Challenges and realities. In W. W. Tornow (C
ward feedback: The ups and downs of it. Symposium condu

Tenth Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Timmreck, C. W., & Bracken D. W. (1996, April). Multisource assessment: Reinforcing the preferred "means" to the end. In W. W. Tornow (Chair), *360-degree feedback systems: Another look at their uses and impact.* Symposium conducted at the Eleventh Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.

Tornow, W. W. (1993). Perceptions of reality: Is multi-perspective measurement a means or an end? *Human Resource Management, 32,* 221–229.

Van Velsor, E., & Leslie, J. B. (1991a). *Feedback to managers: Vol. 1. A guide to evaluating multi-rater feedback instruments.* Greensboro, NC: Center for Creative Leadership.

Van Velsor, E., & Leslie, J. B. (1991b). *Feedback to managers: Vol. 2. A review and comparison of sixteen multi-rater feedback instruments.* Greensboro, NC: Center for Creative Leadership.

Van Velsor, E., Taylor, S., & Leslie, J. B. (1993). An examination of the relationships among self-perception accuracy, self-awareness, gender, and leader effectiveness. *Human Resource Management, 32,* 249–263.

Walker, A. G. (1997, April). *Upward feedback: Incremental improvement in managers' performance over five years.* Paper presented at the Twelfth Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.

Yammarino, F. J., & Atwater, L. E. (1997, Spring). Do managers see themselves as others see them? Implications of self-other rating agreement for human resources management. *Organizational Dynamics,* pp. 35–44.

Yukl, G., & Lepsinger, R. (1995). How to get the most out of 360-degree feedback. *Training, 32,* 45–50.

CHAPTER 9

The Advantages and Pitfalls of Self-Assessment in Organizations

Leanne E. Atwater

Self-assessment, the process whereby individuals evaluate their own performance, skills, or attributes, is used for a variety of purposes in organizations. Included among the uses of self-appraisal are input in the performance appraisal process, a means of identifying developmental needs (in some cases by comparing self- and other perceptions for discrepancies), and input in employment selection decisions. This chapter describes the ways self-assessment is used in organizations, discusses factors that can contribute to more and less accurate self-perceptions and self-appraisals, and—most important—provides recommendations for optimally using self-assessments as a tool for self-development and as input for performance appraisal.

Approximately 5 percent of U.S. companies use some type of self-evaluation as part of their performance appraisal process (Ivancevich, 1995). In one application, self-ratings are used as input for the supervisor doing the evaluation as an additional source of information. For example, a number of law enforcement agencies incorporate self-assessment into appraisal on some dimensions because officers typically work alone, without observation by their supervisor (Love & Hughes, 1994). Self-appraisal can

Note: The author thanks Priscilla Cartier for her assistance in preparing this chapter.