

**Incongruity in 360-Degree Feedback Ratings and Competing Managerial Values:  
Evidence from a Public Agency Setting\***

Shahidul Hassan\*  
John Rohrbaugh

Rockefeller College of Public Affairs and Policy  
University at Albany, State University of New York

\*Corresponding author, Email: [russ hassan@gmail.com](mailto:russ hassan@gmail.com), Phone: (518) 605-6163

Prepared for Presentation at the 9<sup>th</sup> Bi-annual Public Management Research Conference,  
University of Arizona, Tucson

# **Incongruity in 360-Degree Feedback Ratings and Competing Managerial Values: Evidence from a Public Agency Setting**

Shahidul Hassan  
John Rohrbaugh

Rockefeller College of Public Affairs and Policy  
University at Albany, State University of New York

## **Introduction**

An important management technique that has attracted both widespread application and close academic scrutiny is multi-source feedback or 360-degree feedback (Borman, 1997; Waldman, Atwater, & Antonioni, 1998; Toegel & Conger, 2003; Levy & Williams, 2004). This assessment process involves evaluation of managerial effectiveness from four or more sources: supervisors, subordinates, peers, self and possibly others such as customers, suppliers, and consultants (Furnham & Stringfield, 1998). An immediate objective behind the considerable use of 360-degree feedback is enhanced managerial effectiveness, while a long-term goal is improved organizational effectiveness (Furnham & Stringfield, 1998; Levy & Williams, 2004).

Providing feedback to managers about the views of their supervisors, peers, subordinates, and possibly customers potentially offers helpful insights about their strengths and weaknesses and could prompt behavioral change (Mount, Judge, Scullen, Sytsma, & Hezlett, 1998; Waldman et al., 1998). However, to properly guide behavioral change, it is necessary to have somewhat consistent information about managers' performance. The widespread application of 360-degree feedback in performance appraisal has been associated with an attendant interest among researchers in issues of rating congruence. Prior research has extensively studied both self-other<sup>1</sup> and other-other<sup>2</sup> rating congruence and has found rating incongruence among different rater sources, especially between self and other ratings (Harris & Schaubroeck, 1988; Atwater & Yammarino, 1992; Wohlers, Hall, & London, 1993; Furnham & Stringfield, 1994; Atwater, Roush, & Fischthal, 1995; Atwater & Yammarino, 1997; Conway & Huffcutt, 1997; Atwater, Ostroff, Yammarino, & Fleenor, 1998; Furnham & Stringfield, 1998; Atwater, Waldman, Robie, Ostroff, & Johnson, 2005). For this reason, both practitioners and researchers have probed the validity and usefulness of 360-degree feedback for both performance appraisal and human resource development.

Current thinking suggests that rating incongruence exists for valid reasons and should be well integrated in any process of performance appraisal (Tornow, 1993; Borman, 1997; Salam, Cox, & Sims, 1997; Hooijberg & Choi, 2000). However, a theoretically integrated empirical

---

<sup>1</sup> The difference between managers' self and their supervisor, peer, and subordinate ratings.

<sup>2</sup> The differences among managers' supervisor, peer, and subordinate ratings.

exploration of the sources of rating incongruence remains limited. This gap provides the impetus for the current study, including the following key research questions: (1) what factors influence raters' assessment of managers' performance that may contribute to rating incongruence; (2) does rating incongruence exist uniformly in different managerial performance domains; and (3) how can a better understanding of rating incongruence improve the use of 360-degree feedback for human resource development in public organizations?

Prior research on rating incongruence has focused primarily on private rather than public organizations (Church, 1997; deLeon & Ewen, 1997; Hooft, Henk, & Minne, 2006). With the increasing use of 360-degree feedback in public organizations, it is imperative to provide specific guidance for its proper application in performance appraisal and human resource development. One of the key goals of this study is to examine the nature, sources, and managerial implications of rating incongruence to guide the use of 360-degree feedback in public organizations. In comparison to private organizations, public organizations tend to favor higher degrees of formalization with clear division of labor, well-defined departmentalization coupled with stable hierarchical arrangements (Perry & Rainey, 1988; Bozeman, B., Reed, & Scott, 1992; Rainey & Bozeman, 2000; Boyne, 2002). Some observers also suggest that differences exist between public and private organizations with respect to the degree to which personnel rules constrain the administration of employee compensation, termination and merit-based promotion (Rainey, 1983; Rainey, Pandey, & Bozeman, 1995; Rainey, 1997; Kurland & Egan, 1999). This study recognizes that these general structural differences are the results of both effective responses to environmental pressures (Child, 1972), and adaptation to institutional norms (DiMaggio & Powell, 1983) and takes these characteristics as a given context for studying incongruence in 360-degree feedback ratings.

### **Prior Research on 360-degree Feedback**

Given the inherent ambiguity of managerial work, feedback plays an important role in guiding managers' behavior and performance (Ashford & Tsui, 1991). Feedback provides managers with an accurate sense of how others perceive their performance and helps them to identify their strengths and weaknesses (Ashford & Tsui, 1991). Feedback also can improve managers' performance by motivating them and directing them to select appropriate performance enhancing strategies. Managers receive feedback in many ways and from many sources. They may solicit feedback directly from others, or they may infer it from informal cues, or feedback may be inherent in a managerial task itself (Ashford & Tsui, 1991). However, feedback from multiple sources is not always readily available.

Multi-source feedback, also commonly known as 360-degree feedback, is a management technique that is designed to provide managers with feedback from multiple sources. Whereas

traditional performance evaluation methods involve evaluation from a single source, typically supervisors, 360-degree feedback involves evaluation from supervisors, peers, subordinates, and possibly customers in addition to managers' own performance evaluation (Bracken, Timmreck, & Church, 2001). The reasoning behind this approach is that those who have frequent interactions with managers, who are fairly knowledgeable about their performance, and whose opinions are highly valued by them are in the best position to provide them with information that they need to improve their performance (Mount et al., 1998). This line of thinking also is characterized by a shift away from traditional hierarchical structure to matrix structure and team-based organization of work where feedback from multiples sources is considered critical to improve managers' effectiveness.

Aside from the use of multiple raters, 360-degree feedback differs from traditional performance appraisal methods in two ways (Bracken et al., 2001). First, 360-degree feedback is frequently used for human resource development and growth rather than to assist with administrative decisions such as succession planning, promotion, and compensation (Mount et al., 1998; Waldman et al., 1998; Bracken et al., 2001). Second, ratings in 360-degree feedback usually are made anonymously, with the exception of immediate supervisor and not accompanied by face to face discussion (Mount et al., 1998; Bracken et al., 2001). These measures are taken to increase the likelihood that raters provide feedback that is more candid and, hence, more beneficial to managers, as well as to ensure that managers can effectively use feedback from various sources in a non-threatening setting.

Research on 360-degree feedback mainly has focused on the psychometric properties of multi-source ratings, especially on the nature, sources, and practical implications of rating incongruence. Rating incongruence is defined as the degree to which ratings from multiple sources are dissimilar to each other (Mersman & Donaldson, 2000). A high degree of incongruence between self-other ratings generally has been considered undesirable. This is because, to effectively guide behavioral change, it is necessary for managers to receive focused and precise information about the areas in which they need to improve their performance (Borman, 1997). Managers with inconsistent patterns of feedback from their supervisors, peers, and subordinates may find the feedback confusing, unhelpful, and may feel rather unmotivated to rely on it for improving their performance (Miller & Cardy, 2002).

Both self-other and other-other rating incongruence have been studied extensively. Harris and Schaubroeck (1988) conducted a meta-analysis of 36 studies. After correcting for measurement errors, they found a moderate degree of incongruence between self-peer and self-supervisor ratings. They also found that the degree of incongruence between self-other ratings was higher for managerial or professional employees than that for blue-collar employees. In addition, they found a lower degree of incongruence between supervisor-peer ratings than that

between self-other ratings, and incongruence between self-other ratings was particularly high for managerial and professional employees. In addition, Furnham and Stringfield (1994) found a moderate degree of incongruence between self-subordinate ratings across 30 managerial behaviors but a very high degree of incongruence between self-supervisor ratings on the same items. Conway and Huffcutt (1997), in another meta-analysis, found higher degrees of incongruence between self-supervisor, self-peer, self-subordinate ratings and a moderate degree of incongruence between supervisor-peer ratings. Similar results have been found in more recent studies (Atkins & Wood, 2002; Atwater et al., 2005; Hooft et al., 2006). All of these studies indicate that managers tend to have different view of their performance than the views held by others. Further, there seems to be some consensus that supervisor, peer, and subordinate ratings are more closely related to each other than they are to managers' self ratings. Therefore, the following two hypotheses guided the present study:

**Hypothesis 1:** There will be a high degree of incongruence between self-other ratings.

**Hypothesis 2:** The degree of incongruence between self-other ratings will be higher than that between other-other ratings.

The lack of convergence between self and other ratings has led researchers to probe the sources of rating incongruence. Previous research indicates that self-perception (Ashford, 1989) plays an important role in managers' self-evaluation of their performance, and it is influenced by a variety of personality, cognitive, and social factors (Yammarino & Atwater, 1993; Warr & Bourne, 1999). For example, level of self-esteem can lead to either overestimating or underestimating managers' own performance (Ashford, 1989). Managers high in self-awareness are found to better incorporate information from comparisons of behaviors than those low in self-awareness (Nasby, 1989; Wohlers & London, 1989; Yammarino & Atwater, 1993). Additionally, various egocentric biases and cognitive factors can alter managers' self-perception. Self-ratings are known to suffer from leniency (Fox & Dinur, 1988; Fahr & Dobbins, 1989), social desirability (Podsakoff & Organ, 1986), and attributional biases (DeVader, Bateson, & Lord, 1986).

As performance appraisal can be highly threatening experience for many individuals, cognitive factors such as self-protection (Wohlers & London, 1989), self-concept maintenance (Atwater & Yammarino, 1992), and self-presentation pressures (Ashford, 1989) can lead to unconsciously altering self-ratings in an effort to protect a positive self-image. These problems are particularly evident when 360-degree feedback ratings are used for evaluative purposes rather than for managerial development and growth (Bracken et al., 2001). This observation is important for academic research as developmental type of 360-degree feedback initiatives, when properly conducted, should yield less biased self-ratings (Waldman et al., 1998).

Nevertheless, prior research indicates that managers, in general, tend to rate themselves significantly higher than their observers for a variety of egocentric biases and cognitive factors (Harris & Schaubroeck, 1988; Atwater & Yammarino, 1992; Furnham & Stringfield, 1994; Church, 1997; Atwater et al., 2005; Hooft et al., 2006). Based on this evidence, a third hypothesis was suggested:

**Hypothesis 3:** Managers will rate themselves higher than their supervisors, peers, and subordinates.

A great deal of research has investigated the performance implications rating incongruence. Atwater and Yammarino (1992) examined both predictors and outcomes of self-other agreement and found self-other agreement was a moderator of predictor-leadership and performance-leadership relationships. One of the key findings of their study was that correlations between leader behavior and performance were highest for the leaders who were in agreement with their subordinates' ratings as opposed to those who were either overraters or underraters. Van Velsor, Taylor, and Leslie (1993) found that managers who overrated their performance received the lowest subordinate ratings compared to managers who underrated themselves or were in agreement with their subordinates' ratings. Similarly, Furnham and Stringfield (1994) found managers who were rated by their supervisors as more successful tended to have fewer discrepancies between their self and subordinates' ratings than those who were rated as less successful. Recently, Feinberg, Ostroff, and Burke (2005) examined 360-degree feedback data for 68 managers in a financial firm and found that leaders who exhibited higher levels of appropriate leadership behaviors were more likely to have followers who agreed in their perception of the leader. Taken together, these studies indicated that self-other ratings tend to be more congruent for high performing managers.

Further, Atwater and Yammarino (1997) suggested that both the degree and direction of self-other rating incongruence are relevant to managers' individual and organizational performance outcomes. They predicted overraters to be poor performers and less effective managers, as they may misdiagnose their strengths and weaknesses and fail to see the need for further training and development (Mersman & Donaldson, 2000). They also suggested that the effectiveness of underraters would be mixed. This group of managers can be good performers but, since they exhibit low self-confidence, they may have difficulty in making accurate decisions about their future career goals and objectives.

Empirical evidence indicates that performance of overraters tends to be lower than those who are underraters and in agreement with their supervisors, peers, and subordinates (Bass & Yammarino, 1991; Atwater & Yammarino, 1992; Van Velsor et al., 1993; Furnham & Stringfield, 1994). However, it is not necessary that overraters always remain overraters.

Several studies (Atwater et al., 1995; Smither, London, Vasilopoulos, Reilly, Millsap, & Salvemini, 1995; Johnson & Ferstl, 1999; Bailey & Fletcher, 2002) have shown that when overraters receive poor ratings from others, they are able to adjust both their ratings and performance over time. These findings are very important as they show positive longitudinal impact of multi-source feedback on improving managerial effectiveness. For the underraters, the impact of feedback on performance over time remains inconclusive. Both Atwater and others (1995) and Johnson and Ferstl (1999) found underraters to be good performers and that they increased their self-ratings after receiving feedback from others. Yet, while Atwater and others (1995) found no change in their subsequent performance indicated by their subordinates' ratings, Johnson and Ferstl (1999) found their performance decreased.

Although rating incongruence is traditionally considered undesirable, current thinking suggests that it exists for valid reasons and should be integrated effectively in the performance appraisal process. Bozeman (1997) and others (Tornow, 1993; Murphy & Cleveland, 1995; Borman, 1997; Salam et al., 1997; Hooijberg & Choi, 2000) have suggested that rating incongruence may not be a problem because raters from different organizational levels may evaluate different aspects of a manager's performance or simply use different information in their evaluations. Multiple perspectives, even when apparently contradictory, are not necessarily an error but provide insightful information (Tornow, 1993; Salam et al., 1997). They also have argued that the utility of multi-source feedback actually derives from understanding the nature of rating differences observed across rater levels and interpreting them accurately to guide managers' behavioral change (Borman, 1997; Mersman & Donaldson, 2000). Therefore, one should not automatically expect nor desire a high degree of rating congruence (Borman, 1997; Mersman & Donaldson, 2000). In fact, a high degree of rating congruence may be undesirable as it would indicate that additional ratings only offer redundant information and collecting them is a waste of organizational resources (Borman, 1997; Mersman & Donaldson, 2000).

Several studies provide empirical support for such reasoning. While examining 360-degree feedback data for 456 professionals in a public agency, Pulakos, Schmitt, and Chan (1996) discovered a significant difference in the fit of the performance effectiveness models of peers and supervisors. Salam and others (1997) examined whether the relationship between leaders' behavior and their subsequent performance would differ depending on the roles of the raters. Their longitudinal field study found that leaders who were seen as challenging the status quo and encouraging subordinates' independent action were rated lower by their supervisors but rated higher by their subordinates. They concluded that the lack of congruence in ratings from different sources was not necessarily an error but stems from legitimate differences in the structural role of the person rating the manager. Hooijberg and Choi (2000) examined the

extent to which raters differed in the leadership roles that they associate with leadership effectiveness. They found that, depending on the rater-ratee relationship, different leadership roles were associated with leadership effectiveness. They concluded that, “if organizational constituents differ fundamentally in the leadership roles they associate with effectiveness, 360-degree feedback researchers should not only pay attention to rater agreement, but also to substantive differences” (p. 342). Thus, a fourth hypothesis was suggested following the results of these studies.

**Hypothesis 4:** The degree of incongruence between self–other ratings will vary in different performance domains.

### **A Competing Values Approach to Managerial Effectiveness**

To clarify the sources of rating incongruence, Borman (1997) summarized the relevance of perspectives of raters. In essence, a rater may sample different managerial behavior, employ different weight system, and use different schemes of interpretation. A factor that appears to be critical in this perspective is the managerial roles that each rater would consider important and relevant. Managers assume a wide range of responsibilities associated with their position. Those who observe and evaluate them, including the managers themselves, may have different perspectives about what managerial attributes are considered important and relevant.

Vandenberg, Lance and Taylor (1997) stated that “our personal conceptualization of what constitutes performance is largely a function our social perceptions of the individual being rated, and our interaction goals with that individual due to our position in the organization” (p.29). From this perspective, the relative organizational position of the raters will have a significant bearing on the managerial attributes they would select, emphasize, and evaluate. Several conceptions regarding managerial jobs and roles are available. Fayol (1930) provided the now classic description of managerial functions including planning, leading, organizing, staffing and controlling. Katz (1974) identified technical, human, and conceptual skills and their relative importance for effective managers. Mintzberg (1973) concluded from his innovative study that managers perform ten different roles that include figure head, leader, liaison, monitor, disseminator, spokesman, entrepreneur, disturbance handler, resource allocator, and negotiator. This study clearly suggested that managerial work is highly complex. Many factors contribute to this complexity, including the variety of roles (often conflicting) that managers have to perform in their jobs, the sheer number of responsibilities that managers have, and the number of constituents and stakeholders they need to attend every day. Quinn and Rohrbaugh’s (1983) Competing Values Approach (CVA) to organizational effectiveness provides a theoretical framework to articulate the complexity of various managerial roles and helps to better understand the sources of rating incongruence in 360-degree feedback.

The CVA was originally proposed and applied to study organizational effectiveness (Quinn & Rohrbaugh, 1981; Quinn & Cameron, 1983). Later works based on the CVA proceeded beyond the organizational level of analysis and has been applied to study performance at both the group (Reegan & Rohrbaugh, 1990; McCartt & Rohrbaugh, 1995) and individual levels (Denison, Hooijberg, & Quinn, 1995; Quinn, Faerman, Thompson, & McGrath, 2003). The CVA suggests that the criteria for organizational effectiveness can be organized according to three value dimensions. The first value dimension is related to organizational focus, from an internal and micro emphasis on the well-being and development of people in the organization to an external, macro emphasis on the well-being and development of the organization itself. The second value dimension is related to organizational structure, from an emphasis on stability to an emphasis on flexibility. The third value dimension is related to organizational means and ends, from an emphasis on critical processes to an emphasis on organizational outcomes. The relationships between these three sets of value dimensions and the nine effectiveness criteria are shown in Figure 1.

These three sets of competing values are recognized tensions in organizational life. The first pair of competing values, flexibility as opposed to stability, reflects a basic organizational conflict. Diverging viewpoints in considering organizational control and stability versus innovation and change are central to most debates in organizational behavior and theory. The second pair of competing values, internal versus external, represents another basic dilemma in organizational life. From an external perspective, the organization is a rationally designed tool for accomplishing its tasks. The emphasis here is on maintaining organizational viability through acquiring resources and adapting in changing environments. From an internal perspective, the organization is a socio-technical system where participants have unique feelings and preferences and require consideration, appropriate information and stability in their workplace. However, when the external value on the overall organization is maximized, the internal emphasis on the socio-technical balance may be reduced. Similarly, when the emphasis on internal harmony grows, it may shift away emphasis from the overall viability of the organization. Finally, the last pair of competing values, the means-ends dilemma, point to the basic problem of balancing means (processes) and ends (outcomes) in every day organizational life.

One of the key contributions of the CVA is in its connection drawn between the three value dimensions of organizational analysis and Parson's theory of functional prerequisites for any system of action. An orthogonal representation of the first two of the competing values result in four distinct models of organizational analysis in quadrants that match Parson's specification of functional prerequisites: (1) the human relations model, (2) the open systems model, (3) the rational goal model, and (4) the internal process model. The human relations model

emphasizes organizational flexibility, has an internal focus, and stresses on the effectiveness criteria shown in the upper left section of Figure 1: cohesion and moral (as means) and human resource development (as ends). The open systems model emphasizes organizational flexibility, has an external focus, and stresses on the effectiveness criteria as those shown in upper right section of Figure 1: flexibility and readiness (as means) and growth, acquisition, and external support (as means). The rational goal model emphasizes organizational control, has an external focus, and stresses on the effectiveness criteria as those shown in the lower right section of Figure 1: planning and goal setting (as means) and productivity and efficiency (as ends). The internal process model emphasizes organizational control, has an internal focus, and stresses on effectiveness criteria as those shown in the lower left section of figure 1: information management and communication (as means) and stability and control (as ends). This model favors an orderly work situation with appropriate coordination and distribution of information to provide organizational members a sense of continuity and security. The third value dimension, means to ends, is reflected in each of these four models as each model is concerned with both process and outcome effectiveness of an organization.

The notion that raters may sample different behavior, use different cues, and apply different weighting systems (Borman, 1997) in performance evaluation suggests that they may emphasize different attributes of managerial effectiveness due to their unique vantage points and interaction goals with others in the organization. Therefore, managers, as well as their supervisors, peers, and subordinates, are likely, when providing their assessments, to focus on the roles that they believe are most relevant to complete their work activities. Thus, there will be differences in 360-degree feedback ratings to the extent to that they view managerial effectiveness differently and rely upon different frames of reference in their evaluations.

Managers, for example, need to be responsive to the expectations of all who are in their sphere of influence. They are required to perform particular sets of roles that can be somewhat contradictory in nature; yet, to be effective, they need to pay attention to the fulfillment of all of these roles over time. Thus, when evaluating their own performance, managers typically must focus on all of the roles suggested in the four models of the CVA. Subordinates, on the other hand, are interested in gaining guidance and support from their managers to effectively complete their tasks, as well as to achieve personal growth and fulfillment in the organization. These managerial roles are emphasized in the human relations and internal process models of the CVA. When evaluating their managers' performance, subordinates are likely to use a schema that is largely based on the information about the roles reflected in these two models of the CVA. Peers have different subunit goals but must work together to coordinate interrelated activities to accomplish their respective goals. Because these interdependencies are sequential or even reciprocal in nature (Thompson, 1967), peers need to effectively coordinate and control

their tasks and activities. Therefore, the managerial roles that peers are interested in are emphasized in the internal processes and rational goals models of the CVA. Finally, supervisors are charged with the direction of multiple work units. They are mainly interested in the managerial roles that pertain to the flexibility and external focus of the organization. Thus, when they evaluate managers under their leadership, they are likely to use a schema that is largely based on the roles reflected in the rational goals and open systems models of the CVA.

### **Research Method**

A large local government agency worked closely with a research team from a public university located in the south western region of the United States to establish an academy for leadership training. The training program targeted middle-level managers who were eligible to apply for a division chief position in their departments. All of the trainees were required to take a prerequisite test that assessed their preparedness to participate in the program. In addition, each department within the public agency determined the order in which the trainees attended the program based on their job performance, the number of training slots allocated, location of the training sessions, and the needs of the department. The program consisted of a total of 100 hours of classroom instruction over a 13-week period in various offsite training facilities. As part of the training, 360-degree feedback surveys were distributed to provide participants with additional information regarding their leadership potential. At the end of the training program, trainees received 360-degree feedback in the form of individualized reports. They also received one-to-one coaching from the university research team to help them in understanding and making further use of the 360-degree feedback.

The research team at the university had extensive experience in consulting with a variety of public and private organizations on 360-degree feedback initiatives. They used insights gained from those experiences in designing the survey and customizing it according to the unique environment of the public agency. They carefully reviewed the agency's mission, vision, and goals and used the advice of its leaders to select items for the survey. Prior to administering the survey, a pilot initiative was undertaken to ensure that both the survey procedures would be effective and survey items would be both relevant and clear. A group of trainees and their supervisors, subordinates, peers and customers completed the pilot surveys. Later they provided detailed feedback about their experience with the survey procedures and usefulness of 360-degree feedback data to the research team in focus group discussion sessions. Based on the results of the pilot study, survey items were further evaluated and revised for the main initiative.

To collect data, trainees were asked to contact their supervisors, peers, subordinates, and customers and request them to provide a candid assessment of the trainee's performance along

several key managerial behavior and task dimensions. Each trainee also was asked to complete a self-evaluation of personal performance. A total of five cohorts of the leadership trainees were involved in this study. A total of 110 trainees plus their supervisors, one or more peers, one or more subordinates, and in some cases their customers were asked to return completed surveys by mail. All of the 110 trainees completed self-assessments and 68 (62%) received feedback from their supervisors as well as one or more peers and subordinates. Of those 68 trainees, 52% were male and 48% were female.<sup>3</sup> As only a few trainees received customer feedback, those data were not included in this study.

All of the respondents were informed about the intent of the performance evaluation process and how the feedback data would be used constructively in the training program. From the pilot to actual implementation to post hoc interviews, trainees were familiarized with the developmental objectives of the survey by the research team and instructors. Their peers, supervisors, subordinates and customers also received similar communication from both the training academy and the agency about the importance of the survey, as well as its intended use. To ensure confidentiality, all respondents were told that no one other than the trainees would receive a copy of their assessments. Additionally, peers, subordinates, and customers of each trainee were notified that their ratings would not be reported to the trainee unless two or more responses were received from the same rating source. Supervisors, however, did not receive such instruction because each trainee had only one supervisor. With frequent interactions between the research team, training academy, and public organization, program instructors and public agency leaders effectively served as champions for the 360-degree feedback initiative. All of these measures were taken to enhance rater motivation and reduce rating biases (Bracken et al., 2001).

The final version of the survey was printed on four pages. The cover page specified the purpose of the survey, provided an assurance of confidentiality, asked respondents to identify their relationship with the trainee (supervisor, customer, direct report or peer), and reaffirmed that responses would be aggregated for feedback and leadership development only. The second page explained the scales used for measuring the trainee's performance, so that each evaluator was familiar with both the nature of overall scale and each rating. It then highlighted the public agency's statement of its organizational goals of service excellence, fiscal integrity, organizational effectiveness, and workforce excellence. The next two pages presented 48 items designed to measure each rater's satisfaction with the way the trainee being rated undertook certain managerial tasks or exhibited certain behaviors. Each of the 48 questions was phrased initially as: "how satisfied are you with the way this individual...?" Each question was

---

<sup>3</sup> The survey did not include any questions regarding the raters' demographic and other organizational information. Therefore, raters' demographic and organizational characteristics cannot be reported in this paper.

measured using a five-point scale from 1 = dissatisfied to 5 = highly satisfied. In addition, the survey provided space for comments and encouraged respondents to use additional sheets if needed. The two questions eliciting written feedback specifically asked: (1) what the individual being rated did most effectively, and (2) what constructive suggestions could be provided.

Of the 48 items contained in the survey instrument, 12 were identified as being especially central to the four quadrants of the CVA. Thus, the other 36 items were not included in this study. In addition only one survey from a peer and one from survey from subordinate was randomly selected for each trainee. To replace missing values in the data, the following rule was applied. If there was any alternative peer or subordinate evaluation with completed data to substitute for the one with missing data, then the complete data of the alternative peer or subordinate evaluation was used. If no substitute existed, then the missing item code was replaced with the arithmetic mean of the remaining items of the relevant scale. After replacing the missing values, the final data set consisted of a total of 272 surveys (68 from each rating source).

## Results

Table 1 presents means, standard deviations, and bivariate correlation coefficients of the 12 survey items used in this study. Trainees received fairly high performance ratings ( $M = 4.41$ , on a five-point scale) from all of the rating sources. The items in which trainees received the highest and lowest performance ratings were “*listens to ideas and suggestions of those doing work*” ( $M = 4.52$ ) and “*consistently researches new technologies to determine their applicability to service delivery*” ( $M = 4.19$ ), respectively. The standard deviations of all the items were relatively low on average ( $SD = 0.77$ ). The items with the highest and lowest variability in ratings were “*consistently researches new technologies to determine their applicability to service delivery*” ( $SD = .90$ ) and “*insures unit goals are clearly aligned with organization and department’s goals*” ( $SD = .67$ ), respectively. All of the bivariate correlation coefficients were statistically significant ( $p < .05$ ). Further, correlations between items within the CVA quadrants were consistently higher than correlations with items in other quadrants.

Exploratory factor analysis was performed to organize the 12 survey items into the four dimensions of theoretical interest. This process involved identifying survey items that match the CVA framework and then examining their patterns with principal component analysis and varimax rotation. To check for consistency of the resultant model, factor analysis results of the combined data (272 surveys) were compared with the factor analysis results of each separate source of ratings (68 surveys per source). A model with four factors that matched the CVA emerged clearly from this work. These four factors were: (1) supporting people for the human relations quadrant, (2) enhancing agility for the open systems quadrant, (3) providing stability

for the internal processes quadrant and (4) improving results for the rational goals quadrant of the CVA. These four factors together accounted for nearly 80 percent of the rating variance and each of them loaded heavily on three items from the survey.

Table 2 presents results of the principal component analysis. Reliabilities (as shown with Cronbach's  $\alpha$ ) for all of the four factors were high: .84 for supporting people, .91 for enhancing agility, .80 for providing stability, and .88 for improving results. The factor analysis results were highly consistent both in the combined data and in each separate source of ratings. However, factor loadings from the combined data exhibited somewhat simpler structure than factor loadings from separate sources of ratings. The consistency of high factor loadings (on average, .79) for all 12 items suggested that factor scores<sup>4</sup> would be the most useful measure to conduct further analysis of rating incongruence.

Several methods are available to examine rating incongruence; difference score and correlational approaches seem the most intuitive ones. However, neither of these is a reliable method to examine rating incongruence (Mersman & Donaldson, 2000). Prior studies that have used a difference score approach experienced problems in correctly specifying the functional form of the relationships between self-other ratings and performance outcomes (Edwards, 1993a; Edwards, 1993b; Edwards, 1994; Atwater et al., 1998; Johnson & Ferstl, 1999; Mersman & Donaldson, 2000). The correlational approach also is problematic because of the instability of correlation coefficients (Mersman & Donaldson, 2000). This study relied on one-way analysis of variance (ANOVA) with each of the four sets of factor scores as the dependent measure and rating source as the independent measure to examine the patterns of rating incongruence. The four analyses of variance were independent of each other because the factor scores were orthogonal (i.e., the factor scores were uncorrelated). Table 3 presents a summary of the ANOVA results. These results showed that mean performance ratings differed significantly ( $p < .05$ ) in three of the four factors -- enhancing agility ( $F = 7.96$ ), providing stability ( $F = 12.12$ ), improving results ( $F = 7.64$ ) -- by source of ratings. Tukey's HSD multiple comparison tests allowed for testing of all four research hypotheses.

Hypothesis 1 proposed that there would be a high degree of incongruence between self-other ratings. Tukey's HSD multiple comparison tests revealed that in two of the four factors -- enhancing agility and providing stability -- trainee ratings differed significantly ( $p < .05$ ) from all other sources of ratings. In improving results, trainee ratings differed significantly ( $p < .05$ ) from supervisor ratings. There were no significant differences in supporting people. Thus,

---

<sup>4</sup> Factor scores were computed as regression variables in the following way. For each case on a particular pattern, the case's value on each variable was multiplied by the pattern weight for that item. Then, the sum of these weight-times-data products for all the variables yielded the factor score. Factor scores are standardized, which means they have been scaled so that they have a mean of zero and about two-thirds of the values lie between +1.00 and -1.00. Cases will have high or low factor scores as their values are high or low on the variables entering a pattern.

anticipated differences were found on three of the four CVA dimensions. These results provided partial empirical support for hypothesis 1.

Hypothesis 2 proposed that the degree of incongruence between self-other ratings would be higher than the degree of incongruence between other-other ratings. As shown in Table 3, the degree of incongruence between supervisor-peer, supervisor-subordinate, and peer-subordinate ratings was low in three of the four factors -- supporting people, providing stability, and enhancing agility. None of the Tukey's HSD comparison tests was statistically significant in these three factors. In improving results, supervisors did provide significantly ( $p < .05$ ) higher ratings than peers and subordinates. Thus, only two out of 12 possible differences were identified. These results provided clear empirical support for hypothesis 2.

Hypothesis 3 proposed that trainees would provide significantly higher ratings than their supervisors, peers and subordinates. The ANOVA results in Table 3 showed that trainees, on average, provided significantly lower ratings than their supervisors, peers, and subordinates. Further, the mean factor scores of trainee ratings were negative in all of the four performance domains (supporting people =  $-.12$ , enhancing agility =  $-.49$ , providing stability =  $-.58$ , and improving results =  $.31$ ). These results indicated that, instead of overrating, trainees underrated their performance. Thus, hypothesis 3 was not supported by the results of this study.

Hypothesis 4 proposed that the degree incongruence between self-other ratings would vary in different performance domains. To examine variation within each separate source of ratings, one-way repeated measures ANOVA was performed. Table 4 summarizes the results of these four repeated measures ANOVA. Results indicated that trainee and supervisor ratings differed significantly ( $p < .05$ ) across performance domains. Post hoc tests revealed that trainee ratings in providing stability were significantly ( $p < .05$ ) lower than their self ratings in supporting people (providing stability – supporting people =  $.46$ ). Post hoc tests also revealed that supervisor ratings of trainees in improving results were significantly ( $p < .05$ ) higher from their ratings in both providing stability and supporting people (improving results - providing stability =  $.46$  and improving results - supporting people =  $.42$ ). Figure 2 provides a full accounting of the extent of incongruence between trainee-other and other-other ratings across four performance domains. As summarized in Table 3, the incongruence between trainee-other ratings was high and the incongruence between other-other ratings was low in providing stability and enhancing agility due to low trainee self ratings. The significant ( $p < .05$ ) mean rating differences in providing stability were the following: trainee – supervisor =  $-.61$ , trainee – peer =  $-.84$  and trainee – subordinate =  $-.86$ . The significant ( $p < .05$ ) mean rating differences in enhancing agility were the following: trainee – supervisor =  $-.65$ , trainee – peer =  $-.54$ , trainee –

subordinate = -.73. For improving results, the incongruence between trainee-other ratings and the incongruence between other-other ratings was attributable to the high ratings given to the trainees by their supervisors (See Table 4). The significant ( $p < .05$ ) mean rating differences in improving results were the following: trainee – supervisor = -.75, supervisor – peer = .60, and supervisor – subordinate = .45. Finally, the degree of incongruence between trainee-other and other-other ratings was low in supporting people. These findings provided strong empirical support for hypothesis 4. Figure 3 provides a graphical illustration of the variation in the degree of rating incongruence across performance domains in a four-dimensional space.

### **Discussion and Conclusion**

This study provided a basic theoretical framework based on the CVA to organizational effectiveness to explore the nature and sources of incongruence in 360-degree feedback ratings. Statistical results reported in this study provided preliminary evidence in support of the conceptual model, as well as confirmed and extended the literature on rating incongruence. For example, the ANOVA results were consistent with previous findings (Harris & Schaubroeck, 1988; Atwater & Yammarino, 1992; Wohlers et al., 1993; Furnham & Stringfield, 1994; Conway & Huffcutt, 1997) that managers tend to have a somewhat different view of their own performance than the views held by others. Peer and subordinate ratings of managers' performance again were found to be more closely related to each other than to managers' self-ratings. In addition, the present study demonstrated that rating incongruence can vary not only between sources, but also in different performance domains. This is an important finding as it suggests that the same rater (e.g., trainees and their supervisors) may have varying perspectives on different aspects of managerial effectiveness. It also underscores to the fact that one should not automatically expect a high degree of congruence in 360-degree feedback ratings.

The repeated measures ANOVA results showed that rating incongruence in the domain of improving results was high because supervisor ratings were significantly higher than trainees' self, peer, and subordinate ratings. The significantly higher supervisor ratings in improving results might be attributed to the trainee selection process in the management development program. The trainees who participated in this study were a select group of high performing middle level managers. One might speculate that, if trainees had been selected from different levels of the public agency's hierarchy, then a greater degree of variation would have been observed between different sources of ratings. Further, the survey items in the domain of improving results mainly captured the trainees' performance in task-related behaviors. As task-related behaviors are more readily observed and subsequently quantifiable than other

behaviors (Bales, 1950), convergence of trainees' self ratings with their peer and subordinate ratings in improving results seem justifiable.

The repeated measures ANOVA results also indicated that rating incongruence in the domain of supporting people was low because trainees' self ratings were close to their supervisor, peer, and subordinate ratings. The survey items in supporting people tried to capture the trainees' performance in interpersonal behaviors and social aspects of building effective work relationships. To be highly effective, the trainees' performance required a greater degree of shared experience. Thus, convergence of trainees' self ratings with their peer, supervisor and subordinate ratings in supporting people would be expected. However, the results of this study also indicated that rating incongruence in the domains of providing stability and enhancing agility was high, because the trainees provided significantly lower self ratings than their peer, supervisor, and subordinate ratings. This may be attributable to their belief that they are organizationally constrained in fully enacting these managerial roles, a view either not understood or not shared by their supervisors, peers, and subordinates.

Consistent with the findings of Atwater and others (1998) and Johnson and Ferstl (1999), the results of this study indicated that it is important to focus simultaneously on the magnitude and directions (e.g., positive or negative) of managers' self and other ratings to interpret the patterns of incongruence in 360-degree feedback ratings. The variation in incongruence within the same and between different rating sources and in different performance domains found in this study pinpoint to the need for adopting an appropriate statistical method to examine rating incongruence. Had this study used either a difference score or correlational approach to examine rating incongruence, observations concerning the degree and direction of rating incongruence probably would not have been evident.

Previous studies have found that managers tend to provide significantly higher ratings than others due to a variety of cognitive and emotional factors. However, this study found that the trainees typically provided lower self ratings than the ratings of their supervisors, peers, and subordinates. There are several possible explanations for this result. One alternative is that the trainees in this public agency were low in their level of self-awareness and did not accurately diagnose their own strengths and weaknesses. A second and more plausible explanation is that they consciously underestimated most aspects of their performance to avoid an appearance of inflated self-assessment. This could result from uncertainty about the scrutiny that would be given to 360-degree feedback in the training program. Although efforts were taken in the data collection process to reduce such concerns, the trainees in this organization had very little prior experience with 360-degree feedback evaluation methods leading to a modesty bias in the self-rating process.

In an earlier study, Atwater and others (1998) found that the effectiveness of managers was highest when both managers' self and other ratings were fairly high, and when managers' self ratings were substantially lower than other ratings. As indicated by the ratings of their supervisors, peers, and subordinates, the trainees in this study were found to be highly effective even though they underrated their own performance in the domains of enhancing agility and providing stability. Some researchers have suggested that underraters tend to be effective and successful managers, as they have a tendency to overestimate their weaknesses and underestimate their strengths (Ashford, 1989; Atwater et al., 1995; Atwater & Yammarino, 1997; Yammarino & Atwater, 1997; Atwater et al., 1998). Underraters also are thought to be more open to negative feedback than others (Ashford, 1989). The tendency to overestimate their weaknesses perhaps motivates underraters to work hard, in turn may result in greater success at tasks they undertake (Atwater et al., 1998). On the other hand, one also could argue that, because underraters show low self-confidence and self-esteem, they may have difficulty in making important decisions regarding their future career goals and objectives. Nevertheless, as mentioned earlier in this paper, evidence concerning the longitudinal impact of multi-source feedback on the performance of underraters is still inconclusive and requires further empirical research.

There were several limitations in the present study. First, as the survey instrument did not include any question regarding the organizational context and characteristics of the evaluators, this study could not examine if any differences existed in the patterns of rating incongruence by various contextual and demographic factors. Second, also due to of lack of pertinent data, this study could not examine the relationship between rating incongruence, managerial performance and organizational outcomes over time. Thus, an important and interesting future research question is whether underraters change their behaviors and job performance after they receive feedback from their supervisors, peers, and subordinates. Finally, as the data obtained for this study was from a single public agency, further research is needed to determine whether the results of this study have greater generalizability.

One of the most attractive features of 360-degree feedback method is that it gives an initial impression of objectivity and fairness in the performance evaluation process (Toegel & Conger, 2003). When 360-degree feedback is used for performance appraisal, the goal is to minimize rating incongruence and seek a consistent measurement of managers' overall performance (Toegel & Conger, 2003) because it increases the validity of the appraisal process. However, when 360-degree feedback is used for managerial development, the goal is to gather candid information about managers' performance from multiple perspectives, even if these perspectives are contradictory. As shown in this study, the fact that that multiple perspectives within the same or between different sources of ratings and in different performance domains

can lead to rating disagreement is not necessarily a problem but a testament to the reality that organizational actors have different goal expectations. Thus, differences in 360-degree feedback ratings reflect legitimate differences in the raters' perceptions of performance in at least four domains. The efficacy of 360-degree feedback lies in understanding these differences and interpreting them wisely for managerial development to occur.

The increased popularity of multi-source or 360-degree feedback among public sector organizations is a clear demonstration of the desire of public sector managers to improve their performance and increase organizational effectiveness. However, a clear deficiency exists in the literature regarding the contingencies of multi-source feedback use in public agency setting. Considerable pressure has been placed on public agency managers to adopt managerial practices that are popular in the private sector organizations (Golembiewski, 1995; Boyne, 2002). However, care must be taken to evaluate possible differences in work and environmental dynamics and the applicability of research findings obtained from private to public agency settings.

Both researchers and practitioners who apply 360-degree feedback for managerial development in public agencies can draw several useful lessons from this study. First, it is important to consider all sources of input when providing public sector managers with feedback and guidance. Even though this study did not find any difference between peer and subordinate ratings of the trainees, it does not necessarily mean that they are substitutable or unnecessary. Excessive reliance on any particular source of ratings—for example, only supervisors—would be unfruitful and would defeat the main purpose of using multiple rating sources. Second, differences in the perspectives of raters should be considered when providing guidance to managers for their professional development. It is important that public sector managers are aware that those with whom they work may have different goal expectations, and they may provide their evaluations using different priorities for managerial effectiveness. Further, when interpreting 360-feedback ratings to managers, it would be very useful to carefully consider the patterns of congruence and incongruence between their own ratings and the ratings of their supervisors, peers, and subordinates in different performance domains. Insights gained from this information will assist public sector managers to clearly assess their effectiveness with respect to meeting the needs of different stakeholders, as well as to take necessary steps for further performance improvement.

**TABLE 1. MEANS, STANDARD DEVIATIONS AND CORRELATION COEFFICIENTS**

<i>Behaviors</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Supporting People</i>	<i>Enhancing Agility</i>	<i>Providing Stability</i>	<i>Improving Results</i>
Deals effectively with divergent points of view?	4.30	0.79				
Supports people in using their own initiative and creativity?	4.52	0.72	0.64			
Listens to the ideas and suggestions of those doing the work?	4.54	0.70	0.63 0.63			
Stays current with important technical changes in his or her field?	4.36	0.81	0.40 0.43 0.37			
Consistently researches new technologies to determine their applicability to service delivery?	4.19	0.90	0.40 0.42 0.39	0.79		
Participates in training and development programs to strengthen his or her technology skills?	4.45	0.79	0.41 0.40 0.40	0.78 0.74		
Insures unit goals are clearly aligned with organization and department's goals?	4.46	0.67	0.50 0.52 0.47	0.49 0.51 0.52		
Knows and uses all relevant organization's regulations and procedures to achieve unit goals?	4.45	0.71	0.40 0.37 0.40	0.45 0.48 0.42	0.60	
Effectively uses project management techniques to complete goals within established timelines and budgets?	4.25	0.89	0.52 0.50 0.49	0.50 0.50 0.49	0.58 0.57	
Is willing to change existing procedures and practices to improve results?	4.52	0.74	0.45 0.47 0.47	0.48 0.42 0.43	0.50 0.52 0.45	
Acknowledges results achieved as opposed to just following procedures?	4.42	0.70	0.54 0.51 0.47	0.50 0.50 0.49	0.59 0.45 0.53	0.70
Looks for new ways to improve service to customers?	4.45	0.76	0.46 0.56 0.50	0.51 0.49 0.49	0.46 0.46 0.47	0.73 0.73

**TABLE 2. FACTOR PATTERN MATRICES**

<i>Behaviors</i>	<i>Supporting People</i>			<i>Enhancing Agility</i>			<i>Providing Stability</i>			<i>Improving Results</i>		
	All	Trainee Subordinate	Supervisor Peer	All	Trainee Subordinate	Supervisor Peer	All	Trainee Subordinate	Supervisor Peer	All	Trainee Subordinate	Supervisor Peer
Deals effectively with divergent points of view?	<b>0.79</b>	0.31	<b>0.72</b>	0.17	0.04	0.16	0.20	0.32	0.18	0.24	<b>0.74</b>	0.23
Supports people in using their own initiative and creativity?	<b>0.78</b>	<b>0.65</b>	0.29	0.21	0.15	0.16	0.29	0.25	0.23	0.15	0.44	<b>0.72</b>
Listens to the ideas and suggestions of those doing the work?	<b>0.79</b>	<b>0.84</b>	<b>0.83</b>	0.16	0.27	-0.05	0.23	0.24	0.08	0.21	0.13	0.11
Stays current with important technical changes in his or her field?	0.18	0.13	0.01	<b>0.86</b>	<b>0.87</b>	<b>0.87</b>	0.25	0.15	0.12	0.20	0.23	0.19
Consistently researches new technologies to determine their applicability to service delivery?	0.20	0.21	-0.06	<b>0.85</b>	<b>0.82</b>	<b>0.88</b>	0.21	0.29	0.15	0.20	0.19	0.13
Participates in training and development programs to strengthen his or her technology skills?	0.19	0.30	0.22	<b>0.83</b>	<b>0.71</b>	<b>0.86</b>	0.19	0.36	-0.09	0.27	0.24	0.16
Insures unit goals are clearly aligned with organization and department goals?	0.33	0.15	0.47	0.29	0.21	0.04	<b>0.68</b>	<b>0.68</b>	0.32	0.25	0.36	0.52
Knows and uses all relevant agency regulations and procedures to achieve unit goals?	0.11	0.25	0.13	0.21	0.17	0.12	<b>0.83</b>	<b>0.72</b>	<b>0.86</b>	0.28	0.25	0.31
Effectively uses project management techniques to complete goals within established timelines and budgets?	0.40	0.17	0.52	0.30	0.30	0.30	<b>0.65</b>	<b>0.79</b>	0.26	0.17	0.08	0.40
Is willing to change existing procedures and practices to improve results?	0.21	0.23	-0.15	0.18	0.37	0.05	0.29	0.10	0.38	<b>0.82</b>	<b>0.77</b>	<b>0.82</b>
Acknowledges results achieved as opposed to just following procedures?	0.31	0.04	0.41	0.26	0.44	0.13	0.28	0.36	-0.05	<b>0.73</b>	<b>0.71</b>	<b>0.77</b>
Looks for new ways to improve service to customers?	0.31	0.48	0.19	0.28	0.45	0.18	0.14	0.17	-0.14	<b>0.81</b>	<b>0.50</b>	<b>0.84</b>
<b>Cronbach's <math>\alpha</math></b>	<b>0.84</b>	0.36	0.14	<b>0.91</b>	0.18	0.25	<b>0.80</b>	0.14	0.09	<b>0.88</b>	<b>0.85</b>	<b>0.88</b>

**TABLE 3: ONE-WAY ANOVA RESULTS**

<b>Factors</b>	<b>Trainee</b>	<b>Supervisor</b>	<b>Peer</b>	<b>Subordinate</b>	<b>F Ratio</b>
Supporting People	-0.12	0.07	0.21	0.01	1.63
Enhancing Agility	-0.49 <sup>a</sup>	0.15 <sup>b</sup>	0.05 <sup>b</sup>	0.24 <sup>b</sup>	7.96*
Providing Stability	-0.58 <sup>a</sup>	0.02 <sup>b</sup>	0.26 <sup>b</sup>	0.28 <sup>b</sup>	12.12*
Improving Results	-0.31 <sup>a</sup>	0.44 <sup>b</sup>	-0.16 <sup>a</sup>	0.00 <sup>a</sup>	7.64*

Note: Different superscripts indicate significant rating difference at  $p < .05$

\*  $p < .05$ .

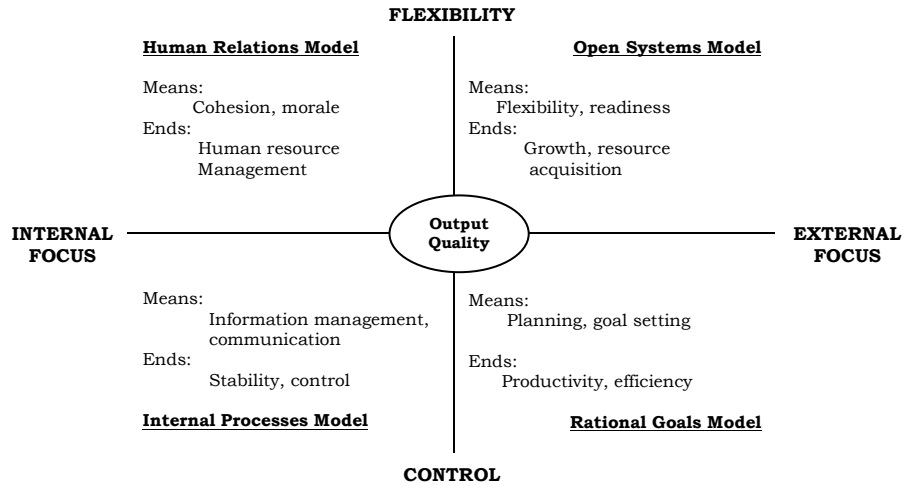
**TABLE 4. ONE-WAY REPEATED MEASURES ANOVA RESULTS**

<b><i>Factors</i></b>	<b><i>Trainee</i></b>	<b><i>Supervisor</i></b>	<b><i>Peer</i></b>	<b><i>Subordinate</i></b>
Supporting People	-0.12 <sup>a</sup>	0.07 <sup>a</sup>	0.21	0.01
Enhancing Agility	-0.49 <sup>ab</sup>	0.15 <sup>ab</sup>	0.05	0.24
Providing Stability	-0.58 <sup>b</sup>	0.02 <sup>a</sup>	0.26	0.28
Improving Results	-0.31 <sup>ab</sup>	0.44 <sup>b</sup>	-0.16	0.00
F ratio	3.67*	2.99*	1.86	1.69

Note: Different superscripts indicate significant rating difference at  $p < .05$

\*  $p < .05$ .

**FIGURE 1. THE COMPETING VALUES APPROACH**

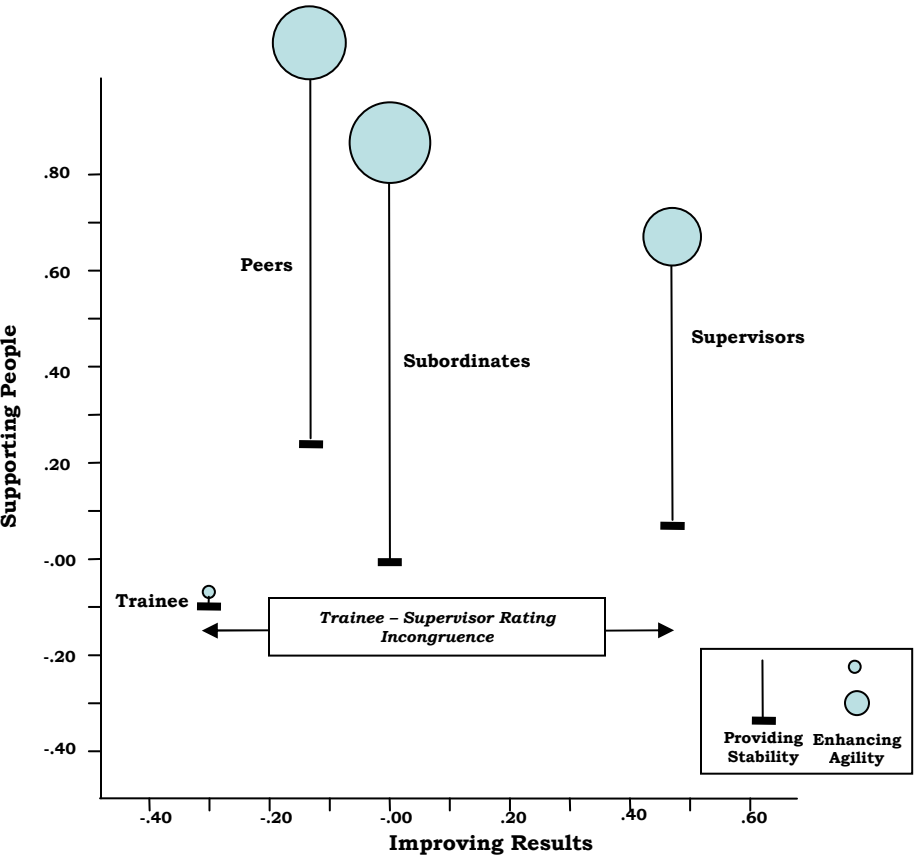


**FIGURE 2: THE DEGREE OF RATING INCONGRUENCE BETWEEN DIFFERENT SOURCE OF RATING  
ACROSS PERFORMANCE DOMAINS**

<b>Trainee and Other Ratings</b>		<b>Other-Other Ratings</b>	
<p><b>Supporting People</b></p> <p>Trainee - Supervisor = -.19 Trainee - Peer = -.33 Trainee - Subordinate = -.13</p>	<p><b>Enhancing agility</b></p> <p>Trainee - Supervisor = <b>-.65</b> Trainee - Peer = <b>-.54</b> Trainee - Subordinate = <b>-.73</b></p>	<p><b>Supporting People</b></p> <p>Supervisor - Peer = -.14 Supervisor - Subordinate = .06 Peer - Subordinate = .20</p>	<p><b>Enhancing Agility</b></p> <p>Supervisor - Peer = .11 Supervisor - Subordinate = -.09 Peer - Subordinate = -.19</p>
<p>Trainee - Supervisor = <b>-.61</b> Trainee - Peer = <b>-.84</b> Trainee - Subordinate = <b>-.86</b></p> <p><b>Providing Stability</b></p>	<p>Trainee - Supervisor = <b>-.75</b> Trainee - Peer = -.15 Trainee - Subordinate = -.31</p> <p><b>Improving Results</b></p>	<p>Supervisor - Peer = -.23 Supervisor - Subordinate = -.25 Peer - Subordinate = -.02</p> <p><b>Providing Stability</b></p>	<p>Supervisor - Peer = <b>.60</b> Supervisor - Subordinate = <b>.45</b> Peer - Subordinate = -.16</p> <p><b>Improving Results</b></p>

Note: Significant mean differences ( $p < .05$ ) are shown in bold

**FIGURE 3: A GRAPHIC ILLUSTRATION OF THE DEGREE OF RATING INCONGRUENCE IN DIFFERENT PERFORMANCE DOMAINS**



## References

- Ashford, S. (1989). Self-assessments in organizations: A literature review and integrated model *Research in Organizational Behavior*, 11, 133-174.
- Ashford, S., & Tsui, A. (1991). Self-Regulation for managerial effectiveness: the role of active feedback seeking. *Academy of Management Journal*, 34(2), 251-280.
- Atkins, P. W., & Wood, R. E. (2002). Self-versus others' ratings as predictors of assessment center ratings: Validation evidence for 360-degree feedback programs. *Personnel Psychology*, 55, 871-904.
- Atwater, L. E., Ostroff, C., Yammarino, F. J., & Fleenor, J. W. (1998). Self-other agreement: Does it really matter? *Personnel Psychology*, 51(3), 577-598.
- Atwater, L. E., Roush, P., & Fischthal, A. (1995). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, 45, 141-164.
- Atwater, L. E., Waldman, D. A., Robie, C., Ostroff, C., & Johnson, J. (2005). Self-other agreement: Comparing its relationship with performance in U.S. and Europe. *International Journal of Selection and Assessment*, 13(1), 25-40.
- Atwater, L. E., & Yammarino, F. J. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, 45(1), 141-164.
- Atwater, L. E., & Yammarino, F. J. (1997). Self-other rating agreement: A review of the model. *Research in Personnel and Human Resource Management*, 15, 121-174.
- Bailey, C., & Fletcher, C. (2002). The impact of multiple source feedback on management development: Findings from a longitudinal study. *Journal of Organizational Behavior*, 23, 853-867.
- Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups*. Chicago, IL: University of Chicago Press.
- Bass, B., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology: An International Review*, 40, 437-454.
- Borman, W. C. (1997). 360 degree ratings: An analysis of assumptions and a research agenda for evaluating their validity. *Human Resource Management Review*, 7(3), 299-315.
- Boyne, G. A. (2002). Public and private management: What's the difference? *Journal of Management Studies*, 39(1), 97-122.
- Bozeman, B., Reed, P., & Scott, P. (1992). Red Tape and Task Delays in Public and Private Organizations. *Administration and Society*, 24(3), 290-322.
- Bozeman, D. (1997). Interrater agreement in multi-source performance appraisal: A commentary. *Journal of Organizational Behavior*, 18, 213-316.

- Bracken, D., Timmreck, C. W., & Church, A. H. (2001). *The Handbook of Multisource Feedback: The Comprehensive Resource for Designing and Implementing MSF Processes*. San Francisco: Jossey-Bass.
- Child, J. (1972). Organization structure, environment, and performance: The role of strategic choice. *Sociology*, 6, 1-22.
- Church, A. H. (1997). Do you see what I see? An exploration of congruence in ratings from multiple perspectives. *Journal of Applied Social Psychology*, 27(11), 883.
- Conway, J. M., & Huffcutt, A. E. (1997). Psychometric properties of multi-source feedback ratings: A meta-analysis of subordinate, supervisor, peer and self-ratings. *Human Performance*, 10(4), 331-360.
- deLeon, L., & Ewen, A. J. (1997). Multi-source performance appraisals. *Review of Public Personnel Administration*, 17(1), 22-36.
- Denison, D., Hooijberg, R., & Quinn, R. E. (1995). Paradox and Performance: Toward a Theory of Behavioral Complexity in Managerial Leadership. *Organization Science*, 6(5), 520-540.
- DeVader, C. L., Bateson, A. G., & Lord, R. G. (1986). Attribution theory: A meta analysis of attributional hypothesis. In E. Locke (Ed.), *Generalizing from Laboratory to Field Settings*. Lexington, MA: Lexington Books.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160.
- Edwards, J. R. (1993a). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal*, 36, 1577-1613.
- Edwards, J. R. (1993b). Problems with the use of profile similarity indices in the study of congruence in organizational research. *Personnel Psychology*, 46, 641-665.
- Edwards, J. R. (1994). The study of congruence in organizational behavior research: Critique and proposed alternative. *Organizational Behavior and Human Decision Processes*, 58, 683-689.
- Fahr, J., & Dobbins, G. (1989). Effects of self-esteem on leniency bias in self-reports of performance. A structural equation model analysis. *Personnel Psychology*, 42, 835-850.
- Fayol, H. (1930). *Industrial and General Administration* (J. A. Coubrough, Trans.). London: Sir I. Pitman & Sons Ltd.
- Feinberg, B., Ostroff, C., & Burke, W. (2005). The role of within-group agreement in understanding transformational leadership. *Journal of Occupational and Organizational Psychology*, 78, 471-488.
- Fox, S., & Dinur, Y. (1988). Validity of self-assessment: A field evaluation. *Personnel Psychology*, 41, 581-592.

- Furnham, A., & Stringfield, P. (1994). Congruence in self and subordinate ratings of managerial practices as a correlate of supervisor evaluation. *Journal of Occupational and Organizational Psychology*, 67, 57-67.
- Furnham, A., & Stringfield, P. (1998). Congruence in job-performance ratings: A study of 360 degree feedback examining self, manager, peers, and consultant ratings. *Human Relations*, 51(4), 517-530.
- Golembiewski, R. T. (1995). *Practical Public Management*. New York: M. Dekker.
- Harris, M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43-61.
- Hooft, E. A., Henk, F., & Minne, M. R. (2006). Construct validity of multi-source performance ratings: An examination of the relationship of self-supervisor-, and peer-ratings with cognitive and personality measures. *International Journal of Selection and Assessment* 14(1), 67-81.
- Hooijberg, R., & Choi, J. (2000). Which leadership roles matter to whom? An examination of rater effects on perceptions of effectiveness. *Leadership Quarterly*, 11(3), 341-364.
- Johnson, J., & Ferstl, K. (1999). The effects of interrater and self-other agreement on performance improvement following upward feedback. *Personnel Psychology*, 52, 271-303.
- Katz, R. L. (1974). Skills of an effective administrator. *Harvard Business Review*, 52, 90-102.
- Kurland, N., & Egan, T. (1999). Public v. private perceptions of formalization, outcomes, and justice. *Journal of Public Administration Research and Theory*, 9(3), 437-458.
- Levy, P. E., & Williams, J. E. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management*, 30(6), 881-905.
- McCartt, A. T., & Rohrbaugh, J. (1995). Managerial openness to change and the introduction of GDSS: Explaining initial success and failure in decision conferencing. *Organization Science*, 6, 569-584.
- Mersman, J., & Donaldson, S. (2000). Factors affecting the convergence of self-peer ratings in contextual and task performance. *Human Performance*, 13(3), 299-322.
- Miller, J. S., & Cardy, R. L. (2002). Self-monitoring and performance appraisal: Rating outcomes in project teams. *Journal of Organizational Behavior*, 21, 609-626.
- Mintzberg, H. (1973). *The Nature of Managerial Work*. New York: Harper & Row.
- Mount, M. K., Judge, T. A., Scullen, S. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater and level effects in 360-degree performance ratings. *Personnel Psychology*, 51(3), 557-576.
- Murphy, K. R., & Cleveland, J. N. (1995). *Understanding Performance Appraisal: Social, Organizational, and Goal-based Perspectives*. Thousand Oaks, CA: Sage.
- Nasby, W. (1989). Private self-consciousness, self-awareness, and the reliability self reports. *Journal of Personality and Social Psychology*, 56, 950-967.

- Perry, J. L., & Rainey, H. G. (1988). The Public-private distinction in organization theory: A critique and research strategy. *Academy of Management Review*, 13(2), 182-201.
- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12, 531-544.
- Pulakos, E. D., Schmitt, N., & Chan, D. (1996). Models of job performance ratings: An examination of rater race, rater gender, and rater level effects. *Human Performance*, 9, 103-119.
- Quinn, R. E., & Cameron, K. S. (1983). Organizational life cycles and the shifting criteria of effectiveness: Some preliminary evidence. *Management Science*, 29, 33-51.
- Quinn, R. E., Faerman, S. R., Thompson, M. P., & McGrath, M. R. (2003). *Becoming a Master Manager: A Competency Framework* New York: John Wiley & Sons.
- Quinn, R. E., & Rohrbaugh, J. (1981). A Competing Values Approach to Organizational Effectiveness. *Public Productivity Review*, 5, 122-140.
- Quinn, R. E., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science*, 29(3), 363-377.
- Rainey, H. G. (1983). Public agencies and private firms: Incentive structures, goals, and individual roles. *Administration and Society*, 15(65-105).
- Rainey, H. G. (1997). *Understanding and Managing Public Organizations*: San Francisco: Jossey-Bass.
- Rainey, H. G., & Bozeman, B. (2000). Comparing Public and Private Organizations: Empirical Research and the Power of the A Priori. *Journal of Public Administration Research and Theory*, 10(2), 447-469.
- Rainey, H. G., Pandey, S., & Bozeman, B. (1995). Research note: public and private managers' perceptions of red tape. *Public Administration Review*, 55(6), 567-574.
- Reegan, P., & Rohrbaugh, J. (1990). Group Decision Process Effectiveness. *Group and Organization Management*, 15(1), 20-43.
- Salam, S., Cox, J., & Sims, H. (1997). In the eye of the beholder: How leadership relates to 360-degree performance ratings. *Group and Organization Management*, 22(2), 185-209.
- Smither, J. W., London, M., Vasilopoulos, N. L., Reilly, R. R., Millsap, R. E., & Salvemini, N. (1995). An estimation of the effects of upward feedback program over time. *Personnel Psychology*, 48, 1-34.
- Thompson, J. D. (1967). *Organizations in Action*. New York: McGraw-Hill.
- Toegel, G., & Conger, J. A. (2003). 360-degree assessment: Time for reinvention. *Academy of Management Learning and Education*, 2(3), 297-311.
- Tornow, W. W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management*, 32(2), 221-229.

- Van Velsor, E., Taylor, S., & Leslie, J. (1993). An examination of the relationships among self-perception accuracy, self-awareness, gender, and leader effectiveness. . *Human Resource Management*, 32, 249-264.
- Vandenberg, R. J., Lance, C. E., & Taylor, S. C. (1997). A latent variable approach to rating sources equivalence: Who should provide ratings on organizational citizenship behavior dimensions? Paper presented at the *Academy of Management Conference*, Boston, MA
- Waldman, D. A., Atwater, L. E., & Antonioni, D. (1998). Has 360 feedback gone amok? *Academy of Management Executive*, 12(May), 86-94.
- Warr, P., & Bourne, A. (1999). Factors influencing two types of congruence in multi-rater judgments. *Human Performance*, 12(3/4), 183-210.
- Wohlers, A. J., Hall, M. J., & London, M. (1993). Subordinates rating managers: Organizational and demographic correlates of self/subordinate agreement. *Journal of Occupational and Organizational Psychology*, 66(3), 263-275.
- Wohlers, A. J., & London, M. (1989). Ratings of managerial characteristics: Evaluation difficulty, co-worker agreement, and self-awareness. *Personnel Psychology*, 42, 235-261.
- Yammarino, F. J., & Atwater, L. E. (1993). Understanding self-perception accuracy: Implications for human resource management. *Human Resource Management*, 32, 231-247.
- Yammarino, F. J., & Atwater, L. E. (1997). Do managers see themselves as others see them? Implications of self-other rating agreement for human resources management. *Organizational Dynamics*(Spring), 35-44.