



Department of Management and Information Systems

Working Paper Series

An Exploration of Factors that Influence the Purposeful Adjustment of Performance Ratings within a Multi-Source Feedback Environment

Cathy L. Z. DuBois Kent State University

Gary A. Kustis

O'Brien, Passen & Associates, Inc.

Robert H. Faley

Kent State University

Debra S. Gatton

Tiffin University

Andrew J. Passen

O'Brien, Passen & Associates, Inc.

(An earlier version of this paper was presented at the 12th Annual Conference of the Society for Industrial and Organizational Psychology, Inc., St. Louis, April, 1997)

Working Paper **MIS2007:12**

Do not cite without permission

ABSTRACT

This exploratory study focuses on factors that influence the purposeful adjustment of performance appraisal ratings given as part of a multi-source feedback (MSFB) process. The relative number of factors that influenced raters and the differential impact of these factors across superior, peer, and subordinate rater groups are examined. Superior, peer, and subordinate perceptions of the MSFB process are also examined. Results reveal that the number of factors does not differ significantly across rater groups but the relative influence of the factors does. Additionally, results reveal that perceptions of the MSFB rating process differ significantly across rater groups. Implications for the MSFB rating process are discussed.

An Exploration of Factors that Influence the Purposeful Adjustment of Performance Ratings within a Multi-Source Feedback Environment

Much current research has examined multi-source feedback programs (MSFB; for a review, see London, 1995). Most of this research has focused on issues common to the performance appraisal literature. For example, much of the MSFB research focuses on either the administration of and the benefits derived from the MSFB process or on the consistency of the ratings given as part of the MSFB process across rater groups (superior, peer, self, and subordinate).

An important issue that the general performance appraisal and the MSFB literatures have paid comparatively less attention to is the purposeful adjustment of performance appraisal ratings. The existing research suggests that raters are often unwilling to make accurate ratings for a variety of reasons. Researchers addressed this as early as the 1950's (c.f. Bass, 1956; Glickman, 1955). Ilgen, Barnes-Farrell and McKellin noted in their 1993 review of performance appraisal research:

There are a wide variety of factors that influence the rating recorded on an appraisal form, some of which are related primarily to the rater's best judgment about the behavior of the person that is rated and some of which are related to other factors such as the rater's willingness to give a low rating to the person even if he or she believes the ratee deserves it. Research on the appraisal process, by choice, focuses upon the judgment component of rating (p. 324).

Similarly, Banks and Murphy (1985) note that "raters are rarely motivated to provide accurate appraisals and may, in some cases, be strongly motivated to provide inaccurate appraisals. Research on performance appraisal has consistently failed to distinguish between evaluation and rating behavior" (p. 343). Finally, Longenecker, Sims, and Gioia (1987) report that "earlier research has either missed or glossed over the fact that executives giving appraisals have ulterior motives and purposes that supersede the mundane concern with rater accuracy" (p. 184). Longnecker et al. (1987) further state that "although

research on rater 'error' has traditionally suggested that raters can and do inflate ratings, researchers have typically not accounted for the realities of the appraisal context to explain why these errors occur" (p. 190).

Issues related to the purposeful adjustment of ratings may be particularly germane when the ratings occur as part of a MSFB process. For example, because the MSFB process employs multiple feedback sources (including superiors, peers, subordinates, and selves), there may be systematic group differences in terms of the factors that influence the purposeful adjustment of ratings. Thus, it would not be surprising to find the factors that influence purposeful adjustments vary as the organizational position of the rater relative to the ratee varies.

Research shows there are often significant differences across rater groups in the performance appraisal ratings they give the same ratee (c.f., Harris & Schaubroeck, 1988; Klimoski & London, 1974; London & Smither, 1995). This research points to three main rationales that may account for these differences. First, rater groups may differ in terms of how they view the ratee's job and how it should be done. Second, rater groups may differ in terms of their ability to observe the ratee's performance. Third, rater groups may differ in the level of their experience and expertise with the rating process. Any of these differences alone or in combination may account for the lack of agreement across rater groups reported in the literature.

A fourth rationale might come into play, as well. We suggest that there may be differences across rater groups in the factors that influence purposeful adjustments of ratings. For example, the influence of liking on purposeful adjustments may be more of an influence for peer raters than superiors or subordinates. Or, peers may be more likely than supervisors to adjust their ratings in order not to hurt the feelings of poor performing coworkers. These varying influences across rater groups could contribute to different amounts and types of purposeful adjustment among raters, resulting in ratings that are not consistent across rater groups.

The Current Study

Very little of the research cited above occurred within an MSFB environment, and much of it

focuses on traditional supervisory ratings (e.g., Tziner, Latham, Prince, & Haccoun, 1996). Just about all of this research focuses on either the rater or ratee exclusively.

The current study simultaneously examines rater and ratee-related concerns associated with the frequency of use of purposeful adjustments within a multi-rater feedback environment that includes supervisory, peer, and subordinate ratings. Within this context, the factors that influence adjustments and the associated rater attitudes about the MSFB rating process are examined.

The current study explores three broadly defined research questions. Question one examines whether the number of factors that influence purposeful rater adjustments differs significantly across rater groups (superiors, peers, and subordinates). Question two examines whether there are significant differences in frequency of use across rater groups in terms of the specific factors that influence purposeful adjustments. Finally, question three examines whether significant differences exist across rater groups in terms of their perceptions of the MSFB process that are potentially salient to purposeful rater adjustments.

The current study examines differences across three rater groups on the frequency of use of seven adjustment-related factors suggested by the performance appraisal literature. We present these factors in two broad categories: rater-related factors and ratee-related factors.

Rater-Related Factors

The four adjustment-related factors examined below deal with the direct or indirect impact of ratings on raters themselves. These include rater concerns about: 1) the rater's personal feelings for the ratee, 2) the inconsistent use of the rating scale across raters, 3) the impact on the rater's work environment, and 4) the impact on the rater's relationship with the ratee.

Rater's Personal Feelings for the Ratee. A rater's personal feelings for a ratee can lead to purposeful rating adjustments. For example, DeNisi and Mitchell (1978) report that raters who evaluate friends typically inflate ratings; Dipboye (1985) reports that rating bias can occur because raters dislike the ratee. Thus, it is not surprising that Stone (1973) noted raters report that evaluation is more difficult if they personally like or dislike the ratee. As stated by Cardy and Dobbins (1986), "it

appears that liking, even though irrelevant to the performance rating task, is a dimension that interferes with raters' evaluations of performance (p. 676).”

Consistent Scale Use. Raters make purposeful adjustments so their ratings appear to be consistent with the ratings of other raters. Sometimes this concern is related to the rater’s desire to be consistent with organizational norms. For example, raters may adjust their ratings simply because others within the organization purposefully adjust ratings (Dipboye, 1985; Kozlowski & Morrison, 1996; Longnecker et al. 1987). That is, raters know that if other raters purposefully inflate their ratings, this unfairly penalizes the ratees of those raters who don’t.

Raters also make purposeful adjustments in response to situational pressures to give higher or lower ratings to ratees (Mero & Motowidlo, 1995). These authors report that situational pressures resulted in purposeful adjustments consistent with those pressures, especially when raters were held accountable for their ratings.

Impact on Rater’s Work Environment. Raters make purposeful adjustments because they believe their ratings will have an impact on their own work environment or organizational outcomes. The research by Longnecker et al. (1987) provides many examples of this. For example, these authors report that many of the 60 high-level executives interviewed revealed that they inflated ratings to keep others from learning about problems within their department, to get a problem employee promoted out of their department, or to avoid unpleasant incidents with problem employees.

Fried and Tieg (1995) report that supervisors with high role-conflict are more likely to inflate ratings out of concerns about the impact ratings will have on their work environment. For example, raters with high role-conflict are more likely to inflate ratings to avoid dealing with time-consuming complaints from disgruntled employees.

Finally, anonymity concerns can influence purposeful adjustments related to rater organizational outcomes. Raters make adjustments when they expect their ratings will become know to the ratee, for example, as part of the performance feedback process (Ilgen & Knowlton, 1980; Klimoski and

Inks,1990). As a result, lack of anonymity can lead to ‘tit-for-tat’ rating scenarios, especially in MSFB rating environments. For example, DeNisi and Baum (cited in DeNisi & Mitchell, 1978) reported that some raters adjust ratings upward if they fear that giving negative ratings to their peers will subsequently result in negative peer ratings for themselves, from those same peers. Similarly, Antonioni (1994) reported that subordinates who rate their superior are inclined to give inflated ratings when their ratings are not anonymous because of fear of supervisory retaliation in the form of lower ratings for themselves. Finally, London, Smither and Adsit (1997) point out that making MSFB raters accountable by making their ratings known does not always result in more accurate ratings.

Impact on Rater-Ratee Relationship. Raters purposefully adjust ratings because of concerns about the impact the ratings will have on their relationship with the ratee. For example, it is common for raters to make upward adjustments (Larson, 1986; Fisher, 1979) to either enhance the rater-ratee relationship or avoid interpersonal difficulties with ratees. Fried and Tieg (1995) point out that giving inflated ratings is a means to gain the gratitude and goodwill of subordinates. The executives interviewed by Longnecker et al. (1987) were well aware of the impact that negative ratings had on their ability to “live with subordinates in a day-to-day relationship” (p.185). Finally, because those who receive negative ratings report greater difficulties dealing with their ratings, they are more likely to create troublesome working relationships as a result (Stone, 1973).

Ratee-Related Factors

The three adjustment-related factors examined below deal with the direct or indirect impact of ratings on ratees. These include rater concerns about the impact their ratings will have: 1) on the ratee’s future performance, 2) the ratee’s organizational outcomes, and 3) on the ratee’s feelings about him/herself.

Impact on Ratee’s Future Performance. Raters make purposeful adjustments (either upward or downward) based on concerns that their ratings will influence the ratee’s future performance. For example, Longnecker et al. (1987) noted that executives reported they deflated ratings to shock poorly

performing subordinates into performing better in the future. Moreover, these executives reported that deflating ratings was a way to teach a rebellious subordinate a lesson.

Alternatively, Longenecker et al. (1987) reported that executives inflated their ratings as a way to encourage employees to maintain or improve their performance. For example, raters inflated their ratings when the ratee's performance had improved during the latter part of the appraisal period, or when the ratee had been suffering personal problems that had negatively impacted their performance. Similarly, Larson (1986) noted that raters are reluctant to give negative ratings because of the discouraging effect it may have on a ratee's performance.

Impact on Ratee's Organizational Outcomes. Raters make purposeful adjustments because they believe ratings will positively or negatively impact the ratee's organizational outcomes. For example, Longenecker et al. (1987) note that executives were more likely to adjust ratings when they believed the ratings would significantly impact the ratee's raises and promotions. These authors report that executives inflated ratings to maximize ratee merit increases as well as to protect ratees against low ratings that might haunt them for the rest of their careers if made part of their permanent records. Finally, the fact that ratings collected for administrative purposes tend to be much more lenient than ratings collected for research purposes (Harris, Smith, and Champagne, 1995) also suggests that raters are reticent to give ratings that negatively impact the organizational outcomes of ratees.

Impact on Ratee's Self-Related Feelings. Finally, raters adjust performance ratings because they believe the ratings will impact how ratees feel about themselves. For example, Larson (1986) reports that raters are reluctant to give negative feedback because they don't want to hurt the feelings of ratees. As stated by one subject in the London, Wohlers and Gallagher (1990) study, "I like my boss and would be less likely to give negative feedback if it would hurt" (p.29). Finally, Longenecker et al. (1987) concluded that executives inflate ratings because they feel sorry for a subordinate, and want to "pick up a guy when he was down" (p. 185).

Purposeful adjustments related to ratee feelings are often the result of rater attributions about the

reasons for poor ratee performance. As noted by Klimoski and Inks (1990) "it is felt that feedback to individuals performing poorly for reasons beyond their control would be different than when this performance is clearly due to a lack of effort" (p. 204). Similarly, Ilgen and Knowlton (1980) report that supervisors were more likely to inflate ratings if they believed the ratee's lack of ability rather than lack of effort was the cause of poor performance.

METHOD

Subjects

Survey data were collected in conjunction with a 360 degree performance appraisal process in a manufacturing division of a Fortune 500 company. Subjects included general managers of this operating division and their staffs. Feedback from superiors was provided by upper-level managers, peer raters included managers of various similar levels, and subordinate raters included first-line supervisory and technical employees.

Data

The 360 degree process was contracted to a consulting firm, which coordinated data collection. Performance rating forms and research questionnaires, along with instructions for raters, were mailed together to all raters. Raters were instructed to complete the performance appraisal form before reading the form marked "For research purposes only". The research instructed raters to complete the form and return it to the consulting firm in a separate, enclosed envelope. The instructions clearly noted that rater responses to the research questionnaire would not be shared with either the ratee or company management. A total of 664 performance rating forms were mailed; 574 usable forms were returned. This represents an 86 percent usable response rate.

The research questionnaire consisted of two sets of items. The first set of items is salient to research questions one and two and the second set to research question three. The first set of items was a literature-generated list of seven factors that might have influenced the performance ratings subjects had just made. The first four factors reflect rater-related concerns and the next three factors reflect ratee-

related concerns. Subjects were asked to indicate which of these factors, if any, had influenced their ratings. These factors are listed in Table 1. In the second section subjects were asked several attitudinal questions related to the nature of the ratings and the rating outcomes. These five items are listed in Table 2.

Analyses and Results

Research-Question One. One-way ANOVA was used to examine whether the number of factors that influence purposeful adjustments differed significantly across rater groups (superiors, peers, and subordinates). Table 3 displays the results of the ANOVA. There is no significant effect due to rater group ($F = .51$; $df 2,553$).

Research-Question Two. Contingency tables with chi-square tests of significance were used to examine whether there were significant differences across rater groups in their use of the seven adjustment factors. Table 4 displays the percentage of each rater group that reported a factor had influenced their ratings. An omnibus chi-square test was performed for each adjustment factor. The results shown in Table 4 reveal that significant differences exist across rater groups on all four rater-related factors, but not on any of the three ratee-related factors. Although rater groups did not differ significantly on the ratee-related factors, note that a very large percentage of all rater groups (69% - 91%) reported that the three ratee-related factors influenced the purposeful adjustment of their ratings.

Cell-level chi-square tests (with alpha-level adjusted for multiple tests) for each significant factor were used to isolate the significant differences across rater groups. These results are displayed in the right-hand columns of Table 4. For the “I personally like/dislike the ratee” factor, significantly fewer superiors (47%) made adjustments than both peers (70%) and subordinates (68%). For the “adjustments to make my ratings consistent with how others are likely to use the rating scale” factor, significantly more peers (28%) made adjustments than both superiors (11%) and subordinates (19%). For the “ratings will have an impact on my work environment” factor, significantly more subordinates (62%) made adjustments than peers (51%). For the “ratings will have an impact on my relationship with the ratee” factor, significantly

more superiors (68%) made adjustments than both peers (42%) and subordinates (39%).

Research-Question Three. MANOVA was used to examine whether there were significant rater-group differences in perceptions of the MSFB rating process. Table 5 displays the results of these analyses. Significant group mean-differences were found for two of the five perceptions noted in Table 2: confidence in rating accuracy ($F = 15.85, p < .001$) and weight of my rating relative to others' ratings ($F = 19.43, p < .001$). Table 6 displays the univariate test results by rater group for these two perceptions. Subordinates (4.2) were significantly more confident in the accuracy of their ratings than peers (3.8). Superiors (3.6) believed their ratings would have significantly more weight in the MSFB process than did either peers (3.1) and subordinates (3.1).

DISCUSSION

The results reported above are consistent with earlier research which established that raters operating within a traditional rating environment purposefully adjust their performance ratings. However, the current study demonstrates that within an MSFB environment the number of factors that influence raters to make purposeful adjustments does not differ significantly across superior, peer, and subordinate rating groups. The average number of factors that influenced performance ratings was about 4.5 (out of 7) for all three rater groups.

More importantly, the current study demonstrates that rater groups are differentially influenced by rater-related factors that influence raters to purposefully adjust their performance ratings. The realities of organizational life dictate differential impacts of the ratings for the different groups of raters. For example, dis/liking influenced the ratings of significantly fewer superiors than peers or subordinates. Perhaps affect is more salient to peers, who work more closely with each other than with their superiors. Similarly, liking can be more salient to a subordinate, who only has one superior to deal with, than to the superior, who has many subordinates to deal with.

Peers were significantly more concerned than superiors or subordinates with making their ratings consistent with the ratings of other raters. Perhaps this is also because they work most closely with each

other. Finally, superiors reported significantly more concern than either peers or subordinates about the potential impact of their ratings on their relationship with the ratee. Perhaps this stems from the fact that superiors, being the more powerful member in the relationship, are often seen as more responsible for maintaining the relationship. Furthermore, difficulties in the rater-ratee relationship that interfere with the superior's ability to manage the ratee's performance can become especially problematic when they impede accomplishment of the superior's assigned organizational goals.

The fact there were no significant differences across rater groups in terms of the three ratee-related factors was not surprising. That is, ratees (whether they are superiors, peers, or subordinates) are impacted similarly by the performance rating process, and raters seem to be aware of this. As noted above, a very large percentage of individuals in all rater groups reported that their ratings were influenced by the potential impact on the ratee's future performance and related organizational outcomes, as well as the ratee's feelings about him/herself. This is very likely due to the fact that all raters are also ratees at some time, and as a result are very much aware of and similarly influenced by ratee-related concerns.

With respect to rater's perceptions of the MSFB rating process, there were only two significant differences across rater groups. First, subordinates were significantly more confident than peers (but not superiors) about the accuracy of the ratings they assigned. However, the differences, as reflected in the effect size, were not great. Second, both peer and subordinate rater groups felt their feedback would be given less weight than feedback from other (i.e., superior) raters. This is not particularly surprising because peers and subordinates are simply less powerful than superiors.

Limitations

There are several limitations about the current study that should be noted. First, as is the case with nearly all survey data, the self-report data used in this study could not be corroborated. Second, although rater training can have an effect on the accuracy of ratings (Sanchez & DeLaTorre, 1996), the raters in the current study were not trained and did not receive information about rating accuracy and bias. Third, raters were all from one organization, and the organization's political climate, culture, and appraisal-

related policies may have influenced the type and amount of purposeful adjustment respondents reported.

Implications

Members of all rater groups in the current study reported their ratings were influenced by rater and ratee-related concerns unrelated to the ratee's actual job performance. This is apparently one of the realities of the rating process and is consistent with the research of Longnecker et al. (1987). These authors concluded that politics in performance appraisal are simply part of the performance appraisal process; politics can be managed but not eliminated. This "reality" should be a valuable reminder to researchers and organizations that raters have agendas in addition to rating accuracy, and that ratings have personal and organizational implications for raters and ratees that cannot be ignored. The results of the current study suggest that the politics of appraisal can easily plague the rating process whether that process is more or less traditional.

Unfortunately, the rush to embrace the MFSB process is all too often based on intuitively appealing anecdotal evidence. Moreover, far too many organizations adopt the MFSB rating process in the hope it will be a panacea for the problems that plague the traditional performance appraisal process. Although this is understandable, organizations must consider both how little we actually know about the relative strengths/weaknesses of the MSFB compared to the traditional process and the high costs of implementing and maintaining an MFSB rating process. What the current study suggests is that it may be prudent for employers to take a wait-and-see approach before they adopt the more complex and costly MSFB process.

One potential ramification of the current study deals with the substance and delivery of MSFB-related training. For example, the results reported above about the amount and nature of rater-group related adjustments suggest that many of the typical recommendations associated with training raters who use a traditional process would apply to those using an MSFB process. This would include specific training about the rater errors (e.g., central tendency, halo, leniency) that are typically associated with the use of more traditional rating processes. However, these training programs very likely will have to be

tailored to meet the needs of individual rater groups. That is, because some errors appear to affect some groups more than others, training programs will have to be tailored for the needs of specific rater groups. Because the MSFB process involves several levels of raters (and ultimately a greater number of raters), this will increase the overall costs of training. However, these customized training programs should increase the overall effectiveness of MSFB systems and their value to organizations. Thus, although the costs of training will increase, the overall net utility derived from the MSFB process should increase also.

There are also a host of new problems that the MSFB process promises related to the legal side of performance appraisal. For example, if an employee did not receive a promotion based in part on the input of peers, could those peers be named as co-defendants in a discrimination lawsuit against the company. Or, if peers were to punish co-workers who violate group production-norms by giving them negative feedback, could they be held liable under tort law. Finally, could a vendor who provides feedback as part of the MSFB process be held partially liable for a negative personnel decision that results from the vendor's feedback. The legal issues associated with performance appraisal that may derive from the MSFB process are poorly understood and obviously need further clarification.

Future Research

In more recent research, Borman, White & Dorsey (1995) suggest that how we view accuracy and bias in performance ratings, particularly with respect to interpersonal factors and affect, may require re-evaluation. These authors suggest that relationships between interpersonal variables and performance ratings might not be indicative of rater bias, but instead reflect meaningful elements of the ratee's contextual performance. For example, purposeful adjustments due to liking/disliking (which many researchers consider bias) might actually reflect important elements of performance, such as the ability to get along or to work in a team-oriented environment. This is consistent with the recent research of Varma, DeNisi & Peters (1996) which suggests that work-related affect might result in part from better subordinate performance.

Clearly there is much that remains to be learned about rater accuracy and the purposeful adjustment of performance ratings. Overall, we understand relatively little about their underlying etiology within the traditional performance appraisal environment, and even less within the more complex environment of MSFB. Future research into the nomological network that surrounds purposeful rater adjustments (especially within an MSFB environment) should concentrate on three issues. Each would yield valuable information that could result in better organizational appraisal-related practices.

First, research should examine whether there are significant differences in the degree (and not just frequency) of the purposeful adjustment in ratings across rater groups in traditional and MSFB performance appraisal environments. This would help organizations better understand relative differences across rater groups in terms of what potentially influences each group's ratings. This might also clarify how various adjustment factors contribute to a better understanding of the meaningful elements of employee performance. Overall, insight into the differential impact of adjustments by different rater groups could help organizations develop more effective ways of dealing with the aftereffects of purposeful adjustments that do not meaningfully reflect worker performance.

Second, once researchers better understand the relative differences across rater groups in terms of both the type and amount of purposeful adjustments, organizations can better implement appraisal-related policies to deal with these differences. For example, it may be possible to directly estimate the systematic error included in ratings as a result of purposeful adjustments that do not meaningfully add to an understanding of employee performance. If so, this error could be removed to make the ratings more interpretable. Moreover, better understanding these rater-group differences should help organizations implement more comprehensive rater training-programs that reduce the influence of these types of purposeful adjustments on rater evaluations.

Finally, research should examine potential micro- and macro-level moderators that may influence the nature of the relationship between the use of purposeful adjustments and rater groups. Some potential micro-level moderators include, among others, the demographic characteristics of the rater and ratee relative to one another, and especially relative to the demographic characteristics associated with the composition of the rater-group. For example, it would be interesting to know whether MSFB-related peer ratings are influenced by the gender composition of the rater group, especially relative to the gender of the ratee. Some potential macro moderators include, among others, organizational factors such as climate, culture, and appraisal-related policies. This also could include whether the firm uses a traditional or non-traditional (e.g., MSFB) approach to the rating process. This information will help organizations better target their limited resources as they attempt to deal with the problems inherent in performance appraisal that compromise the use of appraisal-related information.

Because so many important human resource-related decisions are based on the results of appraisal, much competitive advantage can be gained by better understanding the factors that influence

the purposeful adjustment of appraisal ratings. Organizations have much to gain from insight regarding the factors that influence the outcomes associated with the appraisal process.

REFERENCES

- Antonioni, D. (1994). The effects of feedback accountability on upward appraisal ratings. Personnel Psychology, *47*, 349-356.
- Banks, C.G. & Murphy, K.R. (1985). Toward narrowing the research-practice gap in performance appraisal. Personnel Psychology, *38*, 335-345.
- Bass, B. M. (1956). Reducing leniency in merit ratings. Personnel Psychology, *9*, 359-369.
- Borman, W. C., White, L. A. & Dorsey, D.W. (1995). Effects of ratee task performance and interpersonal factors on supervisor and peer performance ratings. Journal of Applied Psychology, *80*, 168-177.
- Cardy, R.L. & Dobbins, G.H. (1986). Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance. Journal of Applied Psychology, *71*, 672-678.
- DeNisi, A.S. & Mitchell, J.L. (1978). An analysis of peer ratings as predictors & criterion measures and a proposed new application. Academy of Management Review, *3*, 369-374.
- Dipboye, R.L. (1985). Some neglected variables in research on discrimination in appraisals. Academy of Management Review, *10*, 116-127.
- Fisher, C.D. (1979). Transmission of positive and negative feedback to subordinates: A laboratory investigation. Journal of Applied Psychology, *64*, 533-540.
- Fried, Y & Tieg, R. B. (1995). Supervisors' role conflict and role Ambiguity differential relations with performance ratings of subordinates and the moderating effect of screening ability. Journal of Applied Psychology, *80*, 282-297.
- Glickman, A. S. (1955). Effects of negatively skewed ratings on motivations of the rated. Personnel Psychology, *8*, 39-47.
- Harris, M. M. & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-

supervisor ratings. Personnel Psychology, 41, 43-62.

Harris, M. M., Smith, D. E. & Champagne, D. (1995). A field study of performance appraisal purpose: Research- versus administrative-based ratings. Personnel Psychology, 48, 151-160.

Ilgen, D.R., Barnes-Farrell, J.L., & McKellin, D.B. (1993). Performance appraisal process research in the 1980s: What has it contributed to appraisals in use? Organizational Behavior and Human Decision Processes, 54, 321-368.

Ilgen, D.R. & Knowlton, W.A. (1980). Performance attributional effects on feedback from superiors. Organizational Behavior and Human Performance, 25, 441-456.

Klimoski, R. & Inks, L. (1990). Accountability forces in performance appraisal. Organizational Behavior and Human Decision Processes, 45, 194-208.

Klimoski, R. & London, M. (1974). Role of the rater in performance appraisal. Journal of Applied Psychology, 59, 445-451.

Kozlowski, S. W. J. & Morrison, R. F. (1996). Rater strategies: Capturing intentional deviations in performance ratings. Unpublished manuscript.

Larson, Jr., J.R. (1986). Supervisors' performance feedback to subordinates: The impact of subordinate performance valence and outcome dependence. Organizational Behavior and Human Decision Processes, 37, 391-408.

London, M. (1995). Self and Interpersonal Insight: How people gain understanding of themselves and others in organizations. New York: Oxford University Press.

London, M. & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. Personnel Psychology, 48, 803-839.

London, M., Smither, J. W., & Adsit, D. J. (1997). Accountability: the achilles' heel of multisource feedback. Group and Organization Management, 22, 162-184.

London, M., Wohlers, A.J., & Gallagher, P. (1990). 360 degree feedback surveys: A sources of

feedback to guide management development. Journal of Management Development, 9, 17-31.

Longenecker, C.O., Sims, Jr., H.P., & Gioia, D.A. (1987). Behind the mask: The politics of employee appraisal. Academy of Management Executive, 1,183-193.

Mero, N. P. & Motowidlo, S. J.(1995). Effects of rater accountability on the accuracy and the favorability of performance ratings. Journal of Applied Psychology, 80, 517-524.

Sanchez, J. I. & DeLaTorre, P. (1996). A second look at the relationship between rating and behavioral accuracy in performance appraisal. Journal of Applied Psychology, 81, 3-10.

Stone, T.H. (1973). An examination of sex prevalent assumptions concerning performance appraisal. Public Personnel Management, Nov./Dec., 408-414.

Tziner, A., Latham, G. P., Price, B. S. & Haccoun, R., (1996). Development and validation of a questionnaire for measuring perceived political considerations in performance appraisal. Journal of Organizational Behavior, 17,179-190.

Varma, A., Denisi, A. S. & Peters, L.H. (1996). Interpersonal affect and performance appraisal: A field study. Personnel Psychology, 49, 342-360.

Table 1**Factors that Influenced Rating Adjustments**

We are interested in understanding the extent to which each factor influenced the ratings you provided as part of the 360° feedback process. Assign “points” (from 0 to 100) to the items such that those factors that influenced you the most receive more points and those that influenced you the least receive fewer (or no) points.

1. I personally like/dislike this particular individual.
 2. I needed to make adjustments so that my ratings would be consistent with the way I think others are likely to use the rating scale.
 3. These ratings will have a positive or negative impact on my work environment (i.e., morale, climate, etc.).
 4. These ratings will have a positive or negative impact on my relationship with this individual.
 5. These ratings will have a positive or negative impact on the future performance of this individual.
 6. These ratings will have a positive or negative impact on the organizational outcomes for this individual (i.e., promotion, merit increase, termination, etc.).
 7. These ratings will have a positive or negative impact on this individual’s feelings about him/herself.
 8. Other factor(s) which influenced my ratings – please describe in detail.
-

Table 2Perceptions of the MSFB Process Survey Questions

1. Overall, how would you describe your rating of this person?
(from 1, very negative, to 5, very positive)
 2. How surprised do you think the person you rated will be with his/her feedback?
(from 1, not surprised at all, to 5, very surprised)
 3. How confident are you that your rating is an accurate reflection of this individual?
(from 1, not confident at all, to 5, very confident)
 4. As one of a number of people providing ratings for this person, how much weight do you believe your feedback will have?
(from 1, no weight at all, to 5, very great weight)
 5. How much do you believe this person will change due to this feedback process?
(from 1, no change at all, to 5, very great change)
-

Table 3ANOVA: Total number of factors influencing purposeful adjustment by rater group

Effect	df	F
Rater Group	2	.51
Error	553	

<u>Group</u>	<u>Means</u>	<u>Std Dev</u>
Superiors	4.46	1.72
Peers	4.53	1.90
Subordinates	4.36	1.83

Table 4Contingency Tables: Factors that influence purposeful adjustment by rater group

Factor	% of Rater Group that Selected Factor				Cell Level X^2			
	Superior	Peer	Subordinate	X^2	Omnibus	Test Values		
	(1)	(2)	(3)			<u>1-2</u>	<u>1-3</u>	
<u>Rater-related factors:</u>								
Like/dislike of ratee	47	70	68	13.8***	13.0***	10.1***	.2	
Ratings consistent w/others	11	28	19	12.5**	9.6**	2.6	5.5*	
Work environment impact	57	51	62	6.1*	.8	.7	6.1*	
Relationship with ratee	68	42	39	19.4***	15.8***	17.9***	.3	
<u>Ratee related factors:</u>								
Future ratee performance impact	91	85	83	2.3				
Organizational outcomes for ratee	73	74	69	1.5				
Ratee's feelings about self	72	77	69	4.2				

Total N = 556; Superior N = 74; Peer N = 272; Subordinate N = 210.

* p < .05, ** p < .01, *** p < .001

Table 5
ANOVA Analyses: Perceptions of the MSFB Process by Rater Group

Perception	df	F
Positivity / negativity of the rating	2	1.98
Ratee surprise with the rating feedback	2	1.36
Confidence that rating is accurate	2	15.85**
Weight of my rating, relative to others' ratings	2	19.43**
How much the ratee will change due to MSFB	2	2.64

** P < .01

Table 6Univariate Tests for MSFB Process Perceptions: Mean Differences by Rater Group

Attitude	Total	Sup (1)	Peer (2)	Subord (3)	Significant	
Contrasts					<u>1-2</u>	
	<u>1-3</u>	<u>2-3</u>				
Confidence that rating is accurate	<u>M</u>	4.0	4.0	3.8	4.2	
*	<u>SD</u>	.84	.75	.92	.70	
Weight of my rating, relative to others' ratings	<u>M</u>	3.1	3.6	3.1	3.1	* *
	<u>SD</u>	.71	.74	.64	.74	

* p < .01