DISSERTATION

THE INFLUENCE OF INDIVIDUAL AND CONTEXTUAL VARIABLES ON THE RATING BEHAVIORS OF SELF, PEER, SUBORDINATE, AND SUPERVISOR RATERS

Submitted by

Melinda E. Kerst

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2000

COLORADO STATE UNIVERSITY

September 15th, 2000

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY MELINDA E. KERST ENTITLED THE INFLUENCE OF INDIVIDUAL AND CONTEXTUAL VARIABLES ON THE RATING BEHAVIORS OF SELF, PEER, SUBORDINATE, AND SUPERVISOR RATERS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Adviser

Department Head

ABSTRACT OF DISSERTATION

THE INFLUENCE OF INDIVIDUAL AND CONTEXTUAL VARIABLES ON THE RATING BEHAVIORS OF SELF, PEER, SUBORDINATE, AND SUPERVISOR RATERS

Data collected from undergraduate Psychology students who engaged in a performance evaluation task were used to examine the relationships among rater role (self-rater, peer-rater, subordinate-rater, supervisor-rater), rating purpose (administrative, developmental) and rating behavior (rating leniency, discrimination among performance dimensions, discrimination among ratees). Significant main effects were found for rater role on rating leniency and on discrimination among performance dimensions. A significant main effect was also found for rating purpose on rating leniency. In addition, results for an interaction between role and purpose on behavior were significant for rating leniency and for discrimination among performance dimensions.

Individual (self-esteem, locus of control) and contextual (affect, self-efficacy, politics, organizational commitment) variables were examined as covariates of rating behaviors associated with the different rater roles. Although substantial variance in rating behavior was accounted for by these covariates (e.g., self-esteem, locus of control,

affect, efficacy, politics, organizational commitment), results were not in the direction predicted. Individual characteristics and contextual variables were found to account for significant variance in the rating behavior of self-rater. Implications of findings in terms of contributions to the literature and application to organizations are discussed.

Melinda E. Kerst
Department of Psychology
Colorado State University
Fort Collins, CO 80523
Fall 2000

ACKNOWLEDGEMENTS

Thank you to my family who have supported this endeavor with their hearts, with their time, and with unceasing patience. Thank you to my mother and my father for the hours spent cuddling, rocking, strolling, reading to, playing with, and caring for my babies so that I could devote my time to this effort. Without them to provide a loving and steady presence for my children I would not have been able to give this project the attention it required. Thank you also to my parents for loving me exactly as I am. I have always known that with or without a Ph.D., they could not be prouder of me. Their unwavering support has and will continue to inspire me to be the best person I can be both personally and professionally.

Thank you to my husband and number one fan, Frank. He has been at my side from the onset of my career as a graduate student and has without fail supported me in my efforts as a scholar, as a wife, and as a mother. As I complete this project, I know he is beaming with pride!

Thank you to my children, Annabelle and Walker. Whenever this project challenged me, it was their infectious grins and tender hands that provided a welcome retreat. It is with them that I am at my best, and I am constantly amazed at the rewards I receive from the moments spent with them.

Thank you to my sister and brother, Amy and Chad. They have a gift of humor that never fails to engage me. They are my heartfelt supporters and I cherish each moment we spend together.

Thank you to my committee members, Professor George Thornton, Professor Paul Bell, and Professor Michael Slater. I appreciate their insights that have guided this project, and the patience they displayed as this project twisted and turned its way along. Their assistance has contributed greatly to the quality of this project.

Finally, I thank my adviser, Professor Jan Cleveland. Throughout my graduate career, Jan has supported my efforts with patience, with friendship, and with professional guidance of the highest quality. I credit her with encouraging me to strive for and attain goals that I might otherwise have not attempted. Her insights are always motivating, and I welcome the opportunity to benefit from her guidance well into the future.

To Annabelle and Walker

who teach me every day about what really matters in the world

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Traditionally, the term "rater" in the context of performance evaluation has referred to a supervisor giving performance ratings to a subordinate. More recently, the term has broadened to include individuals from all levels within an organization who may potentially provide performance ratings to a variety of ratees. The increasing popularity of processes such as 360-degree feedback illustrates the demand put on organizational members to fulfill multiple rater roles. It has been estimated that some form of multi-rater evaluation is used by 90 percent of Fortune 1000 firms (Atwater & Waldman, 1998).

Raters may typically participate in a multi-rater process in four primary roles: self-rater, peer-rater, supervisory rater (supervisor rating a subordinate), and subordinate-rater (subordinate rating a supervisor; London & Smither, 1995). In programs such as 360-degree feedback, raters participate in all roles as a part of the evaluation process. In other multi-rater programs, raters may only participate in the process in a limited number of roles (for example, as a peer-rater and as a self-rater). With the expansion of performance evaluation beyond that of a supervisor who gives ratings to a subordinate, the potential benefits and complexities of such a process have also increased.

Through the inclusion of additional sources in the process, a variety of individual and organizational benefits may be expected. The feedback process may result in

increased employee participation, better communication between rating sources, and a desire in feedback recipients to seek future performance feedback (Atwater & Yammarino, 1993). These benefits may make evaluation processes that include raters in multiple roles especially valuable to organizations seeking to create highly involved cultures. Indeed, one reason organizations cite for implementing 360-degree feedback and similar processes is as an impetus for culture change (Waldman & Atwater, 1998). While culture change is one use for 360-degree feedback and other multi-rater systems, individual development is the reason most organizations cite for implementing such processes (Waldman & Atwater, 1998).

Evaluation systems in which raters serve in multiple roles are potentially beneficial for both rater and ratee. One benefit of multi-rater evaluation is that raters are given the chance to give feedback to a variety of co-workers. Individuals who are given the opportunity to contribute in this manner may have better working relationships with ratees and may be more committed in their work (Atwater & Yammarino, 1998). On the recipient side of the process, one benefit to ratees is the opportunity to have information regarding their performance contributed to by diverse sources. Ultimately, the goal of multi-rater systems is to provide the recipient of ratings a more holistic view of their performance than would be available through the inclusion of ratings from only one source. The benefit of such systems assumes that ratings given are accurate reflections of performance. Ratings that are discrepant from one another may be limited in their ability

to translate into behavior change in the recipient. With an increase in the number and types of ratings being given, the issue of the quality of ratings coming from raters in different roles becomes important.

Previous research has found that differences may be expected in performance ratings depending on the source of rating. Much attention has been focused on the differences between self and other rating sources (e.g. Harris & Schaubroeck, 1988). Specifically, ratings of one's own performance are generally more lenient when compared to those from other sources. Past research has primarily compared an individual's ratings of self with those received from others. As raters come to fill multiple rating roles in the organization, it is also of value to examine how ratings of self compare with those given to others. Research has not examined the rating behaviors associated with different rater roles. In this respect, the current study makes a unique contribution by examining the quality of ratings that raters give when fulfilling different roles in the rating process. Specifically, the current study examines rating behaviors associated with raters in the role of self-rater and compares them to the rating behaviors associated with the role of other-rater (peer-rater, subordinate-rater, and supervisory rater). In addition, individual and contextual variables are examined for their influence on the rating behavior of raters participating in the aforementioned roles. In the current study, a model of rating behavior is presented in which rater role moderates the relationship between individual and contextual characteristics and rating behavior.

Rater Roles

Rater roles are defined by the relationship of the rater to the ratee. Of interest in the current study are the rater roles of self-rater, peer-rater, subordinate-rater, and supervisor-rater. In the role of self-rater, the rater gives performance ratings to his or her self. In the role of peer-rater, the rater is a co-worker giving performance ratings to peers. In the role of subordinate-rater, the rater is a subordinate giving performance ratings to a supervisor. In the role of supervisor-rater, the rater is a supervisor giving performance ratings to subordinates. These roles have been discussed in the literature as the most commonly held in processes that include multiple raters (Tornow, 1993). Each rater role may be discussed in terms of access they may have to performance information.

Access to information. Murphy and Cleveland (1991) propose that individuals from different organizational levels may differ in the access they have to task behaviors and results and interpersonal behaviors and results. Depending on the rater's relationship to the ratee, they may have a unique opportunity to observe different types of behavior. Depending on whether a rater is in the role of self-rater, peer-rater, subordinate-rater, or supervisor-rater, the type of performance information they may have access to may differ.

When in the role of self-rater, the rater may be viewed as having the most access to information on their performance and results of their performance of any rater role. Specifically, an individual's rating of their own performance may reflect the continuous

access they have to task and interpersonal behaviors and frequent access to the results of such behavior (Murphy & Cleveland, 1991).

When in the role of peer-rater, the rater benefits from being at the same organizational level as the ratee. Ratings given to peers may reflect frequent interaction between rater and ratee. Raters in the role of evaluating a peer may have frequent access to task and interpersonal behaviors and results of behaviors (Murphy & Cleveland, 1991).

When in the role of subordinate-rater (a subordinate rating a supervisor), the rater may benefit from their experience with the interpersonal behaviors of their supervisor. While they may have frequent access to interpersonal behaviors and results of such behaviors, they may rarely have access to task behaviors and only occasional access to the results of such behavior (Murphy & Cleveland, 1991).

When in the role of supervisor-rater (supervisor rating a subordinate), raters may benefit from frequent access to the results of subordinate task behavior. In addition, they may have occasional access to task behaviors, and interpersonal behaviors and results (Murphy & Cleveland, 1991).

The type and level of performance information a rater has about a ratee may differ based on the rater's role in relationship to the ratee. The type of ratings that come from the evaluation process may also differ based on the role of the rater.

Rater Roles and Rating Behavior

As discussed previously, the roles raters take on in processes that involve multiple

raters are defined by the relationship of the rater to the ratee. Findings from previous research suggest that each rater role may be associated with different rating behaviors. The most dramatic differences in rating behavior have been found between those of self-raters and those of other-raters.

Prevalence of Discrepancies between Rating Behaviors of Self-raters and Other-rater

Research indicates that discrepancies between ratings given to self and ratings given to others are common. Reviews of peer, supervisor, and self ratings (e.g., Harris & Schaubroeck, 1988; Mabe & West, 1982) have concluded that, on average, lower correlations exist between ratings from self-raters and those received from others (i.e., peer-raters, supervisory raters) than between other rating sources. Harris and Schaubroeck (1988) found a relatively high correlation between ratings from peer-raters and ratings from supervisor-raters (supervisor rating a subordinate; $r = .62$), but only moderate correlations between self – supervisor ratings ($r = .35$) and self-peer ratings ($r = .36$). Previous reviews have also found self ratings to be associated with greater leniency, less variability, and less discriminant validity than is associated with peer, supervisor, or subordinate ratings (Thornton, 1980).

It is important to note that previous research examining ratings from self and other-raters have based comparisons on differences between self ratings and those received from other sources. In the current study, ratings of self were compared to

ratings given to others. It was expected that regardless of whether ratings of others were examined at the level of given or received, similar biases and reasons for such biases would exist. The prevalence of self- and other- rating disagreement warrants further discussion of the types of rating behaviors each role may engage in.

Self and Other Rating Behaviors

The pattern of discrepancies between ratings of self and ratings of others reported in research may be attributed to the tendency of self and other rating sources to commit different types of rater errors to differing degrees.

Rater errors are typically discussed in terms of three types: how high or low ratings are (leniency), the amount of discrimination within or between ratings of performance dimensions (halo), and the degree to which ratings discriminant between ratees (range restriction). While these three facets of rating behavior may be viewed as contributing to a general impression of how accurate a rating is, they are not direct measures of rating accuracy. Of the three rating errors, the commonly encountered is rating inflation (Landy & Farr, 1983; Longenecker, Sims, & Gioia, 1987).

Rating behavior and self-rater role. Self-other rating disagreement is common and well documented (e.g., Harris & Schaubroeck, 1988; Mabe & West, 1982) The tendency of self-raters to inflate their ratings to a greater degree in comparison to other rating sources has been discussed as a primary reason for self-other disagreement (Steel & Ovalle, 1984). While the current study examines the relationship between

ratings of self and ratings given to other, the same pattern of leniency in ratings of self is expected. Self-raters consistently rate themselves higher than does any other source (Steel & Ovalle, 1984). Self-rating inflation may be a result of ratees' motivation to present themselves in the best possible light. While self-raters may be especially prone to committing leniency errors, this does not mean that the ratings associated with other roles are free from bias.

Rating behavior and other-rater roles. While the ratings given to sources other than self are expected to be more similar than any one source with self, differences may also exist between ratings associated with the roles of peer-rater, subordinate-rater (subordinate rating a supervisor), and supervisor-rater (supervisor rating a subordinate).

Ratings given to peers by co-workers (peer-rater role) have been shown to be valid and reliable (Landy & Farr, 1983), to correlate highly with ratings given to subordinates (Harris & Schaubroeck, 1988), and to be discrepant from self-ratings (Harris & Schaubroeck, 1988). Biases may be especially prevalent when appraisal is conducted for administrative purposes. Employees have been shown to be less receptive to peer appraisal when conducted for administrative as opposed to developmental purposes (McEvoy & Buller, 1987) and have been found to associate more negative outcomes with peer appraisal for administrative as opposed to developmental purposes (Bettenhausen & Fedor, 1997).

Subordinate ratings given to a supervisor (subordinate-rater role) have been found to have a high level of predictive validity in terms of future managerial performance (McEvoy & Beatty, 1989). Ratings of supervisors have been found to be of the same caliber in terms of predictive validity as cognitive ability tests, biodata, and assessment centers (McEvoy & Beatty, 1989). Subordinate ratings of supervisors have been found to be subject to less halo than self-ratings (Holzbach, 1978). In terms of convergent validity and leniency, subordinate ratings of supervisors have been found to be more similar to supervisory ratings of subordinates than self-ratings (Mount, 1984).

Supervisory ratings given to subordinates (supervisor-rater role) are often the standard against which the ratings of others are compared, and are often considered the most accurate of all sources in terms of representing a measure of true performance. However, supervisory ratings are not free from error. Research has provided evidence for the influence of ratee in-group status (Vecchio & Gobdel, 1984), liking of the ratee (Longenecker, Sims, & Gioia, 1987), and supervisor belief system (Jolly, Reynolds, Slocum, 1988) on supervisory rating errors.

While differences may be expected in rating behavior depending on whether the rater is in the role of peer-rater, subordinate-rater, or supervisor-rater, none of the differences in behavior between these roles is expected to be as significant as when ratings are associated with the role of self-rater. Previous research has supported the prevalence of such differences when comparing self ratings to those received from others.

The current study provides a unique contribution by comparing ratings of self with those given to others. Despite this difference in how self-other ratings are to be compared, a similar pattern of differences that was found in previous research is expected in the current study.

Explanations of Self-Other Rating Discrepancies

Research has pointed to a variety of potential reasons for discrepancies between ratings of self and ratings of other. Some researchers (e.g., Klimoski & London, 1974) have proposed that raters from different levels in the organization may have varying perceptions as to what dimensions of performance are most important resulting in rating discrepancies. Other researchers have pointed to differences in opportunities to observe the performance of ratees as a cause of rating discrepancies (Latham & Wexley, 1982).

An additional theory assumes that ratings of self are in some way egocentrically biased. The most common form of egocentric bias is the tendency of self-raters to inflate their ratings in order to better their evaluation (Steel & Ovalle, 1984). An additional form of egocentric bias is taken from attribution theory. Based on his theory, self-raters may attribute good performance to their own behavior, while attributing poor performance to factors external to themselves. Conversely, raters other than self (for example, peer, subordinate, or supervisor raters) may attribute a ratee's good performance to external factors and poor performance to individuals' own behaviors. These differing perspectives may contribute to self-other rating discrepancies.

In a test of the three aforementioned theories, Harris and Schaubroeck (1988) found results of their review to be consistent with an egocentric bias explanation, but not supportive of observational or organizational level explanations.

The context in which feedback processes reside may also contribute to rating discrepancies between those given to self and those given to others. In a study of the influence of context on rating behavior, Cleveland, Murphy, and Tziner (2000) found that raters' beliefs about their organization's performance appraisal system and raters' orientation to the performance system accounted for significant variance in rating behavior. Contextual variables may influence raters in the role of providing ratings of self differently than when raters are in the role of providing ratings to others. Differences in how contextual variables affect rating behavior may contribute to the discrepancy between ratings given to self and those given to others. Research has yet to address this issue.

While previous research has provided valuable information on potential factors that may contribute to the level of agreement between self and other rating sources, this research has not systematically examined antecedents of rating behavior in a situation in which ratings given to self are compared with ratings given to others. Recently, London and Smither (1995) appealed to researchers to examine organizational and personal correlates of agreement between self and other sources. The current study seeks to further research in this area by examining individual characteristics and elements of the

feedback context for their influence on the ratings given to self, ratings given to peers, supervisory ratings given to subordinates, and subordinate ratings given to supervisor.

Current Study

While research has found that ratings of self are often discrepant from ratings received from other sources, it has not systematically examined how ratings of self differ from ratings given to others. A primary goal of the current study is to examine such differences. In addition, little research has addressed what the antecedents of rating behavior are when raters are in the role of rating self versus the role of rating others. Specifically, individual and contextual variables are examined for their influence on the rating behaviors of raters in the role of self-rater and other-rater (peer-rater, subordinate-rater, and supervisor-rater). A model of antecedents contributing to ratings of self and ratings of others is presented in Figure 1.

Overview of Model

In the model presented in the current study, rater role moderates the relationship between individual characteristics and contextual elements of the rating process and rating behavior. It is proposed that the variables that influence rating behaviors may be different for self-rater and other-raters due to differences in the evaluative process they go through. Depending on the rater role, individual and contextual variables may vary in the significance they have to rating behavior. When ratings are associated with self-rater, both individual characteristics may have primary influence on rating behavior. In

contrast, when ratings are associated with rater other than self, the feedback context may have primary influence on rating behavior. Discussion of the variables of interest in the current study follows a general discussion of the evaluative process as experienced by the rater evaluating self and others.

The Process of Evaluation

Rating behaviors may be a function of the accuracy in rater perceptions of the ratee and the motivation of the rater to accurately rate the ratee's performance. This process may be experienced differently when a rater is evaluating his or her own performance as opposed to when he or she is evaluating the performance of others. Elements of the feedback context may contribute to the judgments made by raters and to the accuracy of ratings that they give. In addition, judgments made when rating self may involve ego-defense mechanisms not engaged when performance judgments of others are made. The judgment process and rating process involved when in the role of self-rater versus the role of other-rater may be discussed in terms of similarities and differences.

Accuracy of Ratee-Perception

In a model of the evaluation process proposed by Murphy and Cleveland (1991), performance ratings result from a comparison process in which observed behaviors are examined against the rater's internal standards of performance to determine how well an individual has performed. This comparison process is proposed to differ when the rater is rating self as opposed to others due to the presence of ego-defense mechanisms present in

self-judgment processes. External standards of performance, the rater's beliefs and values, and the rater's individual theories about work combine to create internal standards of performance (Murphy & Cleveland, 1991). This may be true for all rater roles.

External standards come directly from the organization and are often explicitly stated (e.g., work hours, dress codes, standards of behavior). In addition to external standards, the rater's beliefs about the amount of effort required to do a job and the types of results expected from a job may contribute to internal standards. In addition, the individual's theories on what aspects of a job are important or unimportant may contribute to internal standards. Comparisons between observed behaviors and internal standards form the basis on which ratings are made.

A variety of elements present in the feedback context may affect the accuracy with which comparisons are made. The amount of information the rater has about the ratee's performance, the amount of comparative information (information on how well others are doing) the rater has, and the degree to which the rater likes or dislikes the ratee, may all influence the accuracy with which the ratee is perceived by the rater. This may be true for both ratee-perception (in the case of ratings given to peers, supervisors and subordinates) and self-perception (in the case of ratings given to self). If inaccuracies exist in how the ratee is perceived, they may contribute to inaccuracies in ratings. While ratings of self and ratings given to others (i.e., to peer, to subordinate, and to supervisor) both involve comparison of observed behaviors and internal standards to make

performance judgments, ego-defense mechanisms may interfere with the comparisons made when rating self.

By definition, self-assessment involves processing information about one's self and may be influenced by distortive processes that are not present when others are being evaluated (Ashford, 1989). In self-assessments certain ego-defense mechanisms may come into play that are not engaged in when the rater is in the role of peer-rater, subordinate-rater or supervisor-rater.

Because self-rating involves comparing one's own behavior with an internal standard, raters in the role of self-rater may have a tendency to only let positive information about themselves be subject to judgment. Ego-defense mechanisms may serve to protect the self-perception of individuals concerned with maintaining a positive self-image and may result in inaccurate self-perceptions (Ashford, 1989). In attempting to maintain a positive self-image individuals may seek out only performance information that supports a positive self-view, may attempt to avoid messages that threaten their self-view, or may avoid performance information altogether (Ashford, 1989). If negative feedback is received that threatens the individual, he/she may attribute poor behaviors to sources external to oneself. Seeking out only positive information, avoiding information, or distorting information may lead to judgments about one's behavior that are inaccurate. Inaccurate self-perceptions may contribute to inaccuracy in ratings of self.

Motivation for Accurate Ratings

Once the rater forms a judgment about the performance of the ratee, he/she must then assign ratings that reflect how well he/she believes the ratee has performed. The motivation of raters to provide accurate ratings may be a function of whether they believe accuracy will be rewarded, and the degree to which they associate negative consequences with accurate ratings (Murphy & Cleveland, 1991). Elements present in the rating context have important implications for motivation to provide accurate ratings.

While organizations may stress the importance of accurate rating, rarely are valued rewards tied with rating accuracy (Napier & Latham, 1986). Because of this, the rater may not be highly motivated to provide accurate ratings, especially if he/she believes they may be associated with negative consequences. Inaccurate ratings may result if the rater believes that ratings may deny the ratee valued rewards, if the rater fears ratings may result in reduced motivation in the ratee, or if the rater fears that interpersonal relationships with ratees will be damaged by ratings (Murphy & Cleveland, 1991).

An additional reason for distorting ratings may be to avoid negative reactions. Even when raters believe ratings will be kept anonymous, they may hesitate to give low ratings for fear of the ratee discovering that ratings came from them. This may be of particular concern in situations where the pool of available peer, subordinate, or supervisory raters is small.

Various elements present in the feedback context may influence rating accuracy. The purpose of feedback, self-efficacy for the feedback task, perception regarding political maneuvering in the feedback process, and commitment to the organization may influence the degree to which raters are motivated to provide accurate ratings. This relationship is expected for raters in all roles.

In the current study, specific individual characteristics and contextual factors are examined for their influence on the rating behaviors of raters in the role of self-rater and other-rater (peer-rater, subordinate-rater, and supervisor-rater).

Individual Characteristics

Individual characteristics come from within the individual and may determine how the individual approaches the evaluation process. Self-esteem, intelligence, ability, depression, self-consciousness, and self-monitoring are examples of individual characteristics (Atwater & Yammarino, 1997). Two important characteristics for understanding self-perception accuracy and self-rating accuracy that are examined in the current study are self-esteem and locus of control.

As was discussed previously, judgments made about one's own behavior may trigger defensive mechanisms that are not engaged when judgments are made about the behaviors of others. Two individual characteristics that may drive defense processes in self-raters are self-esteem and locus of control (Harris & Schaubroeck, 1988). Because self-esteem and locus of control may contribute to rating accuracy through accuracy in

self-perception, they may be especially influential to the rating behavior of self-raters as opposed to other-raters.

Contextual Variables

As outlined by Murphy and Cleveland (1991), the appraisal process exists in an organizational context which may influence how the rater approaches the rating process and the subsequent performance ratings that are given to ratees. The context in which the feedback process resides may influence rating behavior when the rater is in the role of self-rater, peer-rater, subordinate-rater (subordinate rating a supervisor), and supervisor-rater (supervisor rating a subordinate).

Proximal and distal variables. Organizational context may be discussed in terms of proximal and distal variables. Proximal variables include those that directly influence the rater (Murphy & Cleveland, 1991). Interactions between coworkers, rater perceptions of the ratee, and perceived consequences of rating are all examples of proximal contextual variables. Distal variables include those elements of the organization that indirectly influence the rating process (Murphy & Cleveland, 1991). The climate, culture, and values of the organization are examples of distal contextual variables.

For purposes of the current study contextual variables are grouped into four categories: beliefs about the ratee, beliefs about the feedback system, orientation to the feedback system, and attitudes toward the organization.

"Beliefs about the ratee" involve the level of familiarity the rater has with the ratee's performance. Of specific interest in the current study is the level of affect the rater has for the ratee.

"Beliefs about the feedback system" involve rater beliefs about how the feedback system works. Of specific interest in the current study are the purpose for which ratings are being given and the level of self-efficacy the rater has for rating performance.

"Orientation to the feedback process" involves the rater's sense of trust in and comfort with the appraisal process. Of specific interest in the current study is the level of political maneuvering that raters believe is present in the feedback process.

"Attitudes toward the organization" involve the rater's perceptions of the organization. Of specific interest in the current study is the level of commitment the rater has for their organization.

Contextual variables are expected to account for significant variance in the rating behaviors associated when raters in the role of rating self and when raters are in the role of rating others. Those elements of the feedback context that most directly impact the rater (i.e.,"beliefs about the ratee", "beliefs about the feedback system", "orientation to the feedback process") are expected to account for more variance in rating behaviors than more distal variables ("attitudes toward the organization"). In a study examining the rating behaviors of supervisors, Cleveland et al. (2000) found that organizational attitudes

accounted for little variance in rating behavior while beliefs about the appraisal system and orientation to the system accounted for substantially more variance.

Individual Characteristics and Self-Other Rating Behaviors

Individual characteristics may be especially important to understanding the rating behaviors associated with the role of self-rater. In the current study, self-esteem and locus of control are examined for their influence on rating behavior.

Self-esteem

Self-esteem may be discussed as a measure between an individual's ideal self and how the individual perceives him/herself (Brook, 1991). In general individuals high in self-esteem view themselves positively while those with low self-esteem tend to view themselves negatively.

Individuals high in self-esteem have been found to give themselves more lenient self-ratings (Farh & Dobbins, 1989; Levy, 1993). This relationship has been found to be more salient when evaluations are done on ambiguous performance dimensions (Farh & Dobbins, 1989; Levy, 1993). In a review of self and other rating sources, Harris and Schaubroeck (1988) found evidence that self-raters' motivation to inflate ratings for purposes of enhancing their evaluation was moderated by self-esteem. Specifically, it is expected that high levels of self-esteem will be related to leniency in self-ratings.

Locus of Control

Locus of control involves the degree to which individuals attribute their behavior to events over which they have no control (external locus of control) or attribute their behavior to events under their control (internal locus of control; Rotter, 1966). Research has found that self-raters with internal locus of control report higher levels of performance than do those with external locus of control (Levy, 1991). This may be the result of a protective mechanism in which those self-raters who believe they determine their performance may not want to give themselves poor ratings because this would reflect badly on them. More lenient self-ratings are expected for individuals scoring high for having an internal locus of control.

Contextual Variables and Rating Behavior

Contextual characteristics may influence the rating behaviors of raters in both the self-rater and other-rater (peer-rater, subordinate-rater, supervisor-rater) role. When the rater role is other, contextual variables may be especially influential in determining rating behavior. Beliefs about the ratee, beliefs about the feedback system, orientation to the feedback system, and attitudes toward the organization are discussed for their relationship to the rating behavior associated with the roles of self-rater and other-rater.

Beliefs about the Ratee

The level to which the rater likes a ratee may influence the ratings given as a self-rater, peer-rater, subordinate-rater and supervisor-rater. In the current study, affect is examined for its relationship to rating behaviors.

Affect. The level of "affect" or liking the rater has for the ratee may be particularly important to understanding rating behaviors. The level of liking between the rater and ratee has been related to a variety of rating errors including leniency, halo, and low interrater agreement (Kingstrom & Mainstone, 1985; Tsui & Barry, 1986).

In general, the more raters like a ratee, the more likely they are to give the ratee positive ratings. In a study of supervisory ratings Kingstrom and Mainstone (1985), found that raters who reported higher levels of friendship with a ratee tended to give such ratees significantly higher performance ratings. In one study, three-fourths of managers interviewed admitted to allowing liking to inflate performance ratings and dislike to lower performance ratings (Longenecker, Sims, & Gioia, 1987). Research has also supported an effect for liking on the level of discrimination among ratings. Affect has been found to be positively related to halo (lack of discrimination among performance dimensions) such that raters with positive affect tended to exhibit positive halo whereas raters with negative affect tended to exhibit negative halo (Tsui &Barry, 1986). There also may be a relationship between self-efficacy and affect. Affect has been shown to be more influential on ratings when performance dimensions were less observable and more

ambiguous (Varma, Denisi, & Peters, 1996). It is expected that positive affect will be related to more lenient, less discriminant ratings.

Beliefs about the Feedback System

The behavior of raters may be affected by feelings about their role in the feedback process and perceptions regarding how appraisal information is used in the organization. Self-efficacy for performance evaluation and the purpose for which ratings are being given are examined in the current study for their relationship to rating behavior.

Self-Efficacy. Self-efficacy involves the degree to which the rater believes he or she has the skills needed to adequately evaluate the ratee's performance (Napier & Latham, 1986). Self-efficacy may be influenced by a rater's ability to observe the ratee's performance, by the level of rater training the rater has had, and the rater's experience with rating. Raters higher in self-efficacy are believed to approach the appraisal process more conscientiously than raters who do not believe they have the skills necessary to perform the task. Recent research has shown that raters with higher levels of self-efficacy were more likely to discriminant between the performance of ratees (Cleveland et al.,2000). The relationship between self-efficacy and rating behavior may differ among rating sources.

While higher levels of self-efficacy may be associated with less lenient and more discriminating rating behaviors in peer, subordinate, and supervisory raters, high levels of self-efficacy among self-raters may translate into more lenient and less discriminating

ratings. Of all rating sources, self-raters have the most frequent access to information on the ratees' task and interpersonal behaviors and the subsequent results of such behaviors (Murphy & Cleveland, 1991). This extensive knowledge of behavior may result in high levels of self-efficacy for self-raters.

In the current study, high levels of self-efficacy may be expected to be related to less leniency and increased discrimination for the other-rater role (peer-rater, subordinate-rater, supervisor-rater). A high level of self-efficacy for self-raters may be expected to be associated with leniency and a lack of discrimination among performance dimensions.

Rating purpose. Traditionally, performance ratings from sources other than supervisors have been used for developmental purposes (i.e., indicating strengths and weaknesses). More recently there has been an increase in the number of organizations using ratings from sources other than supervisors for administrative decisions (i.e., pay and promotion; London & Smithers, 1995). Even when the purpose for evaluation is explicitly stated as being for development, raters may doubt that ratings will have absolutely no influence in administrative decisions. The relationship between use of performance ratings (for developmental or administrative purposes) and subsequent rating behavior has received considerable research attention. In the current study, rating purpose (administrative, developmental) is examined as a between-subjects variable.

When raters believe that appraisals will be used to make promotion

decisions, they are more likely to be lenient in their ratings than when they believe the appraisal is being used to provide feedback on employee strengths and weaknesses (Cleveland, Murphy, & Williams, 1989; Landy & Farr, 1983; Murphy & Cleveland, 1991; Zedeck & Cascio, 1982). In addition, rating for administrative purposes has been associated with less discrimination among performance ratings (Zedeck & Cascio, 1982) and among ratees (Cleveland et al., 2000). Rating for developmental purposes has been associated with increased discrimination among ratings of performance dimensions (Cleveland et al., 2000). When administrative decisions are associated with ratings of self, raters are especially likely to inflate their performance ratings (Campbell & Lee, 1988). Peer ratings conducted for administrative purposes have been found to contain greater halo, to be more lenient, less discriminating, less reliable, and less valid than when conducted for developmental purposes (Farh, Cannella, Bedeian, 1991). Accountability for ratings may influence leniency. When raters evaluate performance for administrative conditions and when they have the expectation that their ratings will be validated, they give less lenient ratings than when they do not expect ratings to be validated (Farh & Werbel, 1986). Increased leniency and decreased discrimination may be expected when raters in all rater roles believe ratings will be used for an administrative as opposed to developmental purpose.

The rating role an individual is in may interact with the purpose for rating to influence rating behavior. Because raters in the role of rating self have the unique

opportunity to control the ratings they receive, they may be especially likely to inflate ratings when ratings are associated with an administrative purpose (Campbell & Lee, 1988). The tendency to inflate ratings when in the role of self-rater may be driven by a desire to avoid any negative consequences that he/she associates with low ratings (Longenecker, Sims, & Gioia, 1987). A primary negative consequence that self-raters may associate with low ratings of self may be having desirable rewards withheld by supervisors. This protective mechanism may result in ratings of self associated with an administrative purpose that are especially high and that are associated with especially low levels of discrimination.

Orientation to the Feedback System

Rating behavior may be influenced by the rater's overall approach to the appraisal process. Perceived politics in the feedback process is examined in the current study for its relationship to rating behavior.

Politics in the appraisal system. Politics in the appraisal system involves the rater's perceptions about the direction and degree to which others bias their performance appraisal. Research has shown that performance ratings may be shaped more by a manager's belief system than by a subordinate's actual work performance (Jolly, Reynolds, & Slocum, 1988). Raters who believe that there is a tendency to distort ratings in their organization are more likely to provide more lenient and less discriminant ratings (Cleveland et al., 2000). Social pressure to bias ratings and concern that ratings

may result in conflict have been related to desire to participate in 360 degree appraisal processes (Westerman & Rosse, 1997). This effect may be especially salient raters in the roles of peer-rater, subordinate-rater, and supervisor-rater.

Attitudes toward the Organization

How raters approach the appraisal process may be influenced by their feelings about the organization in which they work. In the current study, commitment to the organization is examined for its relationship to rating behavior.

Organizational commitment. Organizational commitment involves the extent to which an individual identifies with and is involved in the organization (Porter, Steers, Mowday, & Boulian, 1974). One component of organizational commitment that has particular value for understanding rating behavior is attitudinal commitment. Attitudinal commitment is characterized by strong beliefs and acceptance of the organization's goals and values, and a strong desire to continue to be a member of the organization (Mowday, Porter, & Steers, 1982). One approach to understanding the relationship between commitment and rating behavior is by examining of the relationship between self-efficacy and organizational commitment.

Research has found that raters who tend to be high in attitudinal commitment also tend to be high in self-efficacy (Locke, Frederick, Lee, & Bobko, 1984; Mathieu & Zajac, 1990). In the performance appraisal context, it may be that raters who are highly committed to the organization have acquired the necessary skills and goals needed to

effectively appraise an individual's performance. Given sufficient skills and knowledge, rater self-efficacy for the appraisal task may increase, resulting in less lenient and more discriminant ratings. When discussing raters from different levels in the organization, it may be that the organizational role of the rater may also contribute to feelings of commitment and self-efficacy.

A central feature of attitudinal commitment is acceptance of the organization's goals and values (Mowday, Porter, & Steers, 1982). One way rating behavior may be influenced by organizational commitment is through the degree to which the organization values performance ratings coming from different sources in the organization. Traditionally, organizations have relied on supervisors to provide ratings of ratees (Murphy & Cleveland, 1991). Due to their historical presence in the organization, ratings from a supervisor may be perceived as having more value in the appraisal process than from peer, subordinates, or self. A rater's perception that the organization does not value his or her ratings may result in diminished self-efficacy for the task. If raters in the role of peer-rater, subordinate-rater, or self-rater perceive that the ratings they are providing may not be given adequate weight in the appraisal process they may be less inclined to develop the skills necessary to conduct an adequate appraisal. Higher levels of commitment to the organization would be expected to be related to less lenient, more discriminant ratings among all rater roles.

Research Question

Previous research has found evidence of differences in ratings associated with self and other-raters. This research has focused on comparing self ratings with those received from others. In contrast, the current study examines differences in ratings associated with self-rater and other-rater in the context of ratings given. While research indicates that self and other rating differences are prevalent, little research has examined why such differences may exist. Recent evidence suggests that individual characteristics (Atwater & Yammarino, 1997) and elements present in the rating context (Cleveland et al., 2000) may be important determinants of rating behavior. The current study examined both individual and contextual variables for their influence on rating behavior of self-rater and other-rater.

In the current study, the rater role (self-rater, peer-rater, subordinate-rater, and supervisor-rater) and purpose of rating (administrative, developmental) were examined for their influence on rating behavior. The covariate variables of self-esteem, locus of control, affect, self-efficacy, politics in rating, and organizational commitment were also examined for their relationship to rating behavior.

Whenever appropriate, the relationship between attitudinal variables and rating behavior was assessed using three criteria (dependent measures): (a) rating leniency (the degree to which the rater gave high vs. low ratings to the ratee), (b) the degree to which

the rater discriminated between different dimensions of performance, and (c) the degree to which the rater discriminated between ratees.

Hypotheses

Hypotheses were proposed regarding the relationship between rater role, rating purpose, and rating behavior.

H1: A main effect for rater role on rating behavior was hypothesized. The highest and least discriminant (between performance dimensions) ratings would be found when raters were in the role of self-rater (rating self). Ratings associated with the role of peer-rater, subordinate-rater, and supervisor-rater were each expected to differ significantly from those associated with self-rater.

H2: A main effect for rating purpose on rating behavior was hypothesized. The highest and least discriminant ratings (between ratees and between performance dimensions) would be found in the administrative condition as opposed to the developmental condition for all rater roles.

H3: An interaction between rating purpose and rater role on rating behavior was expected. The highest and least discriminant ratings (in terms of performance dimensions) would result from ratings in the self-rater role done for administrative as opposed to developmental purposes.

Hypotheses four through eight involved the covariate variables of interest. Covariates were examined as both explanatory and control variables as outlined below.

H4: An interaction for rater role and self-esteem on rating leniency was hypothesized. Specifically, it was predicted that raters in the role of self-rater and who were high in self-esteem would give more lenient self-ratings than would raters in other roles who are also high in self-esteem.

H5: An interaction for rater role and locus-of-control on rating leniency was hypothesized. Specifically, it was predicted that raters in the role of self-rater and who had an internal locus-of-control would give more lenient self-ratings than would raters in other roles who also had an internal locus-of-control.

H6: Covariates representing "individual characteristics" (self-esteem and locus-of–control) were hypothesized to account for significantly more variance in leniency in ratings when the role was self-rater than covariates representing "contextual factors" (affect, self-efficacy, rating politics, and organizational commitment).

H7: Individual characteristics (self-esteem, locus-of-control) and contextual factors (affect, self-efficacy, rating politics, organizational commitment) were hypothesized to account for significant variance in the rating behavior of raters other than self (peer-rater, supervisor-rater, subordinate-rater). It was hypothesized that of all the covariates studied, organizational commitment would have the weakest relationship to rating behavior for all other-rater roles.

H8: Controlling for the effects of covariates was hypothesized to result in a less significant relationship between rater role (self-rater, peer-rater, subordinate-rater, supervisor-rater), rating purpose (administrative, developmental) and rating behavior (leniency, discrimination between performance dimensions, discrimination between ratees).

CHAPTER II

METHOD

Participants

Students enrolled in an Introductory Psychology course served as participants in the current study (N=179). The average age of participants was under 20 years. Some 87 men and 92 women participated. A moderate level of previous experience with rating the performance of others (M=4.67, SD= 1.42) and themselves (M=4.9,SD=1.48) was reported by participants. Participation in the study fulfilled a course requirement although participation in this specific study was voluntary. Means and standard deviations for demographic variables appear in Table 1.

Design

This study involved a randomized, mixed design (between-within). Rating purpose (between-subjects factor) consisted of two conditions, ratings linked to an administrative purpose and ratings linked to a developmental purpose. Rater role (within-subjects factor) consisted of four conditions; self-rater, peer-rater, subordinate-rater (subordinate rating a supervisor), and supervisor-rater (supervisor rating a subordinate). The study was not fully crossed, the result being that no single participant experienced all levels of the within-subjects factor.

Procedure

Participants were recruited for participation in the study through the use of an information sheet that outlined the purpose and requirements of the study. Participants interested in taking part in the experiment signed up for a time that the study was in session. Groups of three to five participants engaged in the experiment during each session.

Group assignment. Following a brief explanation of the study, participants were assigned at random to participate in the study as either a group member or as a supervisor. Participants assigned to be group members were given name tags to wear with the letters A, B, C, or D on them. Name tags were used to identify group members to be evaluated. Participants serving as supervisors were given an instruction sheet outlining their role in the task. The experimenter explained the instruction sheet to the supervisor prior to the start of the group task, and outside of the hearing range of group members. The supervisory instruction sheet appears as Appendix A.

Task. Following group assignment, all participants were provided with information on the group task. Supervisors were responsible for describing the group task to group members. Participants were instructed that the task involved idea-generation. Specifically, group members were asked to generate ideas on what constitutes appropriate and inappropriate use of student fees. Participants were instructed that they had 15 minutes to complete the idea generation task. Participants were told that upon completion

of the task they would be responsible for presenting their conclusions to the supervisor and the study experimenter.

Group members were told that supervisors would be responsible for describing the task, observing the activities of the group, selecting group members to record and present information, tracking time remaining, and assisting the group if needed. Participants were asked if they had any questions concerning how the task was to be carried out, were reminded of the time limit, and then were instructed to begin the task.

Distribution of measures. Upon completion of the group task, the experimenter provided instructions and measures to all participants. Measures consisted of rating forms, surveys assessing the covariate variables, and a subject information sheet. Instructions and rating forms for group members appear in Appendix B; instructions and rating forms for supervisors appear in Appendix C. Participants who served as group members completed ratings of other group members (peer-rater), of the supervisor (subordinate-rater), and of themselves (self-rater). Participants who served as supervisors completed ratings of group members (supervisor-rater) and of themselves (self-rater). All participants completed surveys assessing the covariate variables of interest. These included measures of self-esteem (Appendix D), locus-of-control (Appendix E), affect (F), self-efficacy (Appendix G), politics in the rating process (Appendix H), and organizational commitment (Appendix I). In addition, participants completed a subject information sheet (Appendix J). The order in which covariate measures were presented

to participants was randomized to reduce the potential for order effects. To insure anonymity, measures were coded and participants were instructed not to put their names (or others) on any materials.

When all measures were completed and collected, participants were thanked for their participation and were instructed that their participation in the study was over.

Manipulation of the independent variables. Rating purpose was manipulated by assigning subjects to one of two rating conditions, rating linked to an administrative purpose (n=85) or rating linked to a non administrative purpose (n=94). Participants in the administrative purpose condition were told that the supervisor or group member with the best performance would have his or her name entered into a lottery. Level of performance was determined by averaging performance ratings from all sources, for each participant. Participants were told that when all data had been collected, the experimenter would draw a name from the lottery pool and the winner would receive one hundred dollars. Participants in the developmental purpose condition were told that their ratings would be used for providing valuable information to researchers on how individuals perform in groups.

Rater role was manipulated by having participants serve in several rater roles (self-rater, peer-rater, subordinate-rater, and supervisor-rater). Each participant who served as a group member rated the performance of all other group members (peer-rater), rated the performance of the supervisor (subordinate-rater), and rated his or her own

performance (self-rater). Each participant who served as supervisor rated the performance of all subordinates (supervisor-rater) and rated his or her own performance (self-rater).

Covariate Variables

Covariate variables of self-esteem, locus of control, affect, self-efficacy, politics in rating, and organizational commitment were examined for their relationship to rating behavior. In addition to these measures, a demographic measure was included for purposes of collecting information on participants such as age, sex, and experience with rating. This measure appears in Appendix J.

Self-esteem. Self-esteem was measured using the 10-item Rosenberg Self-esteem Inventory (Rosenberg, 1965). Some scale items were reverse-coded during the analysis phase of the study. High scores indicate high self-esteem. Cronbach's alpha for the scale was .80. This measure appears in Appendix D.

Locus of Control. Locus of control was measured using a 20-item modified Work Locus of Control Scale (Gupchup & Wolfgang, 1997). The scale was specifically developed to assess the level of control individuals believe they have on the job. For purposes of the current study, phrasing of items was altered to make the measure applicable to a university setting. Some scale items were reverse-coded during the analysis phase of the study. Higher scores on the scale indicate an internal locus of control. Cronbach's alpha for the scale was .72. This measure appears in Appendix E.

Affect. Affect was measured using four items developed from Tsui and Barry (1986). Cronbach's alpha for the scale was .64. This measure appears in Appendix F.

Self-efficacy for rating. Self-efficacy was measured using eleven items taken from scales developed by Napier and Latham (1986). Items assess the level of competency participants believe they have to evaluate the ratee's performance. Items were modified for purposes of examining supervisory self-efficacy in a previous study (Cleveland et al., under review). Cronbach's alpha for the modified scale was .78. High scores on the scale indicate high levels of self-efficacy. This measure appears in Appendix G.

Politics in rating. Perceptions of politics in rating was assessed using a modified version of the 25-item Political Considerations in Performance Appraisal Questionnaire (Tziner, Prince, & Murphy, 1997). Seven items from the original measure were dropped because they were not appropriate for use in the current study, leaving 17 remaining items. Items assess the degree to which raters believe manipulation of ratings occurs in the organization. Cronbach's alpha for the modified scale was .87. This measure appears in Appendix H.

Organizational commitment. Organizational commitment was measured using a 15-item scale adapted from the Organizational Commitment Questionnaire (Mowaday, Steers, & Porter, 1979). Items assess the level of commitment individuals have for their organization. For purposes of the current study, phrasing of items was altered to make the

measure applicable to a university setting. Cronbach's alpha for the adapted scale was .91. This measure appears in Appendix I.

Dependent Measures

Three rating behaviors (i.e., rating leniency, discrimination among ratee, discrimination among dimensions) were examined. Rating behaviors may be discussed in terms of those of participants who served as group members (participated in experiment as a group member who engaged in idea-generation task) and those who served as supervisors (participated in experiment as a supervisor who provided direction to group members).

Group members. For participants who served as group members, rating behaviors were calculated based on the performance ratings they gave to themselves, to their peers, and to their supervisor. Three measures of rating leniency were obtained for each group member. For each group member, a mean of the performance ratings given to themselves, a mean of the performance ratings given to peers, and the mean of the performance ratings given to the supervisor was calculated. Mean performance ratings were obtained by averaging ratings across the six dimensions represented in the performance measure. Because group members often evaluated the performance of more than one peer, mean performance ratings were averaged across peers to obtain an overall mean performance rating for peers. This resulted in a single measure of peer rating leniency for each group member.

For each group member, three measures of discrimination among performance dimensions were obtained. The standard deviation of the performance dimension means was calculated for each group member's ratings of own performance, ratings of peer performance, and ratings of supervisor performance. Because group members often evaluated the performance of more than one peer, level of discrimination among performance dimensions for ratings of peers was calculated by obtaining an overall mean of peer ratings by dimension. The standard deviation of the dimension means was then calculated. Level of discrimination among ratees was represented by the standard deviation of the mean peer performance ratings.

Supervisors. For participants who served as supervisors, rating behaviors were calculated based on the performance ratings they gave to themselves, and to those they supervised (subordinates). Two measures of rating leniency were obtained for each supervisor. For each supervisor, a mean of the performance ratings given to self and a mean of the performance ratings given to subordinates was calculated. Mean performance ratings were obtained by averaging ratings across the six dimensions represented in the performance measure. Because supervisors evaluated the performance of more than one subordinate, mean performance ratings were averaged across subordinates to obtain an overall mean performance rating. This resulted in a single measure of leniency for ratings given to subordinates for each supervisor.

For each supervisor, two measures of discrimination among performance dimensions were obtained. The standard deviation of the performance dimension means was calculated for each supervisor's ratings of their own performance and ratings of their subordinates' performance. Because supervisors evaluated the performance of more than one subordinate, level of discrimination among performance dimensions for ratings of subordinates was calculated by obtaining an overall mean of subordinate ratings by dimension. The standard deviation of the dimension means was then calculated. Level of discrimination among ratees was represented by the standard deviation of the mean subordinate performance ratings.

The rating form used by participants to evaluate performance was developed for use in the current study. The rating measure was based on a taxonomy of job components proposed by Campbell (1990) and dimensions of group performance proposed by Werner (1994). Dimensions were selected for their successful use in previous research for assessing the performance of individuals working on both individual and group tasks across a variety of jobs (Campbell, 1990; Werner, 1994). Cronbach's alpha for the scale used was .92.

CHAPTER III

RESULTS

Hypotheses one through three involved the influence of rater role (rating given to self or given to other) and rating purpose (rating for an administrative or developmental purpose) on rating behavior (rating leniency, discrimination between performance dimensions, and discrimination between ratees). However, another variable was included in analyses: experimental condition (served in the experiment as a group member or supervisor member). Although no specific hypotheses regarding experimental position were made, it is included in analyses based on its importance in the study's overall design.

In the current study, participants served in either the position of group member or in the position of supervisor member. In the experimental position of group member, participants evaluated their own performance (self-rater role), evaluated the performance of their peers (peer-rater role), and evaluated the performance of their supervisor (subordinate-rater role). In the experimental position of supervisor member, participants evaluated their own performance (self-rater) and the performance of those subordinates they supervised (supervisor-rater role). Because the rater role variable is not fully crossed, examination of results at the level of experimental position was warranted. Therefore, results for Hypotheses one through three are discussed in terms of the full

sample of participants, the group member sample (participants who served as group members in the study), and the supervisor member sample (participants who served as supervisor members in the study).

Using the full sample, rater role was analyzed at two levels: self-rater and other-rater. For the group member sample, rater role consisted of three levels: self-rater, peer-rater, and subordinate-rater (subordinate evaluating a supervisor). For the supervisor sample, rater role consisted of two levels: self-rater and supervisor-rater (supervisor evaluating a subordinate). Means and standard deviations for demographic, covariate, and dependent variables are presented in Table 1. Intercorrelations for demographic, covariate, and dependent variables are presented in Table 2.

Main Effects and Interactions: Results for Full Sample

To examine the general effects of the primary independent variables of interest on rating behavior of the full sample of participants (N=179), a series of 2x2x2 general linear model (GLM) analyses were conducted.

Rating purpose (administrative, developmental) and experimental position (group member, supervisor member) served as between-subjects factors in the analysis. Rater role (self or other) served as a within-subjects repeated factor. In order to obtain a measure of ratings for "other-rater", ratings associated with the role of peer-rater, subordinate-rater, and supervisor-rater, were averaged across rater roles.

Two separate 2x2x2 GLM analyses were conducted on the full sample for Hypotheses one through three. The first analysis examined the effects of experimental position (group member, supervisor member), rating purpose (administrative, developmental), and rater role (self-rater, other-rater) on rating leniency. The second analysis examined the effects of the independent variables on discrimination between performance dimensions. Results are discussed in terms of each hypothesis. Discrimination between ratees could not be analyzed in this analysis, as the roles of self-rater and subordinate-rater (subordinate rating a supervisor) only rated one ratee. Table 3 contains a summary of results (main effects, interactions) for the full sample.

Main effect for rater role. Hypothesis one proposed a main effect for rater role (self-rater, other-rater) on leniency and discrimination between performance dimensions. Means and standard deviations for leniency and dimension discrimination associated with each rater role (self, other) are found in Table 1.

Results from 2x2x2 GLM analysis using leniency as the dependent variable were significant for a rater role main effect $F(1, 175) = 61.22$, $p < .01$, $n^2 = .26$. Consistent with the direction predicted, ratings associated with the self-rater role ($M = 5.95$, $SD = .87$) were significantly more lenient than those associated with the role of other-rater ($M = 5.51$, $SD = .85$).

Results from GLM analysis using discrimination between performance dimensions as the dependent variable were also significant for a rater role main effect,

$F(1, 175) = 14.43$, $p < .01$, $n^2 = .08$. In the direction hypothesized, ratings associated with self-rater were less discriminating between performance dimensions ($M = .42$, $SD = .33$) than were ratings of other-rater ($M = .53$, $SD = .23$).

Main effect for rating purpose. Hypothesis two proposed a main effect for rating purpose on leniency and discrimination between performance dimensions. Means and standard deviations for rating purpose by rater role for the full sample are found in Table 4.

Results from 2x2x2 GLM analysis using leniency as the dependent variable were significant for a rating purpose main effect, $F(1, 175) = 14.92$, $p < .01$, $n^2 = .08$. As hypothesized, ratings associated with an administrative purpose ($M = 5.96$, $SE = .09$) were significantly more lenient than those associated with a developmental purpose ($M = 5.50$, $SE = .08$). Rating purpose was not found to influence discrimination between performance dimensions.

Rater role and rating purpose interaction. Hypothesis three proposed an interaction between rater role and rating purpose on leniency and discrimination between performance dimensions.

Results from 2x2x2 GLM analysis using leniency as the dependent variable were significant for an interaction between rater role and rating purpose on leniency $F(1, 175) = 50.941$, $p < .01$, $n^2 = .23$. Findings from the interaction were based on Estimated Marginal Means (EM) that were generated for purposes of the analysis. Consistent with

the direction predicted, ratings associated with the self-rater role when rating for an administrative purpose (EM= 6.44, SE= .09) were significantly more lenient than those associated with self-rater role when rating for a developmental purpose (EM = 5.52, SE = .09). Ratings associated with other-rater were similar for the administrative (EM = 5.47, SE = .10) and developmental purpose (EM = 5.47, SE =.10). The interaction of rater role and rating purpose on rating leniency for the full sample is shown in Figure 2.

Follow-up t-tests were used to compare interactive effects across self-rater and other-rater. When rating for an administrative purpose, self-raters were significantly more lenient (M = 6.37, SD = .79) than were other-raters (M = 5.48, SD = .80), t (1,84)=11.195, p<.01. When rating for a developmental purpose, no significant difference was found between self-rater (M =5 .56, SD = .76) and other-rater (M = 5.54, SD=.90) in terms of leniency.

Results from GLM analysis using discrimination between performance dimensions as the dependent variable were also significant for an interaction between rating purpose and rater role, F(1, 175) = 9.09, p < .01, η^2 = .05. In the direction hypothesized, ratings associated with the self-rater role when rating for an administrative purpose (EM =.36, SE =.04) were significantly less discriminant (between dimensions) than those associated with self-rater role when rating for an developmental purpose (EM =.47, SE = .04). Ratings associated with other-rater were similar for the administrative (EM = .55, SE = .03) and developmental purpose (EM =.49, SE = .03). The interaction

of rater role and rating purpose on dimension discrimination for the full sample is shown in Figure 3.

Follow-up t-tests were used to compare interactive effects across self-rater and other-rater. When rating for an administrative purpose, self-raters were significantly less discriminant (M =.37, SD =.35) than were other-raters (M = .56, SD = .25), t (1,84) = 5.04, $p < .01$. When rating for a developmental purpose, no significant difference was found between self-rater (M = .46, SD = .30) and other-rater (M = .51, SD = .22) in terms of discrimination.

Additional analyses. Additional analyses examined the relationship between experimental position (participated in experiment as group member or supervisor member) and rating behavior. No significant results were found in terms of main effects or interactions involving experimental position on either leniency or discrimination between performance dimensions.

Main Effects and Interactions: Group Member Sample

A 3x2 GLM analysis was conducted on the group member sample (n=128). Data from one subject was missing from analyses due to failure to complete all rating measures. Rater role (self-rater, peer-rater, subordinate-rater) served as a within-subjects repeated factor, rating purpose (administrative, developmental) served as a between-subjects factor, and rating behaviors (rating leniency, discrimination between performance dimensions, discrimination between ratees) served as the dependent

measures. Due to the design of the study, the rater role of supervisor was not included in these analyses (it is reported later in results).

Two separate 3x2 GLM analyses were conducted on the group sample for Hypotheses one through three. The first analysis examined the effects of rater role and rating purpose on rating leniency. The second analysis examined the effects of the independent variables on discrimination between performance dimensions. In addition, for Hypotheses two, discrimination between ratees served as the dependent measure. For the group sample this refers to peer-raters giving ratings to peer ratees. Results may be discussed in terms of Hypotheses one through three. Table 5 contains a summary of results (main effects, interactions) for the group sample.

Main effect for rater role. Hypothesis one proposed a main effect for rater role (self-rater, peer-rater, subordinate-rater) on leniency and discrimination between performance dimensions. Means and standard deviations for leniency and dimension discrimination associated with each rater role are found in Table 1.

Results from 3x2 GLM analysis using leniency as the dependent variable were significant for a rater role main effect, $F(2, 125) = 17.03$, $p < .01$, $n^2 = .21$. Consistent with the direction hypothesized, results from paired t–tests found ratings of self ($M = 5.93$, $SD = .89$) were significantly higher than those associated with either peer-rater ($M = 5.60$, $SD = .77$) or subordinate-rater ($M = 5.53$, $SD = 1.11$). Also as predicted, ratings

associated with the peer-rater role and those associated with subordinate-rater role did not significantly differ in terms of leniency.

Results from GLM analysis using discrimination between performance dimensions as the dependent variable were also significant for a rater role main effect, $F(2, 125) = 9.15$, $p = .01$, $\eta^2 = .13$. Paired t-tests were conducted to determine where significant differences existed. Consistent with the direction hypothesized, the level of discrimination between performance dimensions associated with self-rater ($M = .42$, $SD = .34$) was significantly less than that associated with either peer-rater ($M = .51$, $SD = .29$) or subordinate-rater ($M = .58$, $SD = .31$). It was hypothesized that no significant differences would exist between levels of dimension discrimination between peer-raters and subordinate-raters. Contrary to this, ratings from peer-raters were associated with significantly less discrimination between performance dimensions than were ratings from subordinate-raters.

Main effect for rating purpose. Hypothesis two proposed a main effect for rating purpose on leniency, on discrimination between performance dimensions, and on discrimination between ratees. Means and standard deviations for rating purpose and rating behaviors by rater role for the group sample appear in Table 6. There was no significant main effect of purpose on any of the three dependent variables.

Rater role and rating purpose interaction. Hypothesis three proposed an interaction between rater role and rating purpose on leniency and discrimination between performance dimensions.

Results from 3x2 GLM analysis on the group member sample using leniency as the dependent variable were significant for an interaction between rater role and rating purpose $F(2, 125) = 21.24$, $p < .01$, $n^2 = .25$. Findings from the interaction were based on Estimated Marginal Means (EM) that were generated for purposes of the analysis.

In the direction hypothesized, the highest ratings were associated with the role of self-rater when rating for an administrative purpose (EM = 6.29, SE = .10) as opposed to a developmental purpose (EM = 5.62, SE = .10). In contrast, ratings associated with peer-raters were slightly higher when done for a developmental purpose (EM = 5.71, SE = .09) as opposed to an administrative purpose (EM = 5.47, SE = .09). Ratings associated with a subordinate-rater role were similar under administrative (EM = 5.54, SE = .14) and developmental (EM = 5.51, SE = .14) conditions. The interaction between rater role and rating purpose on rating leniency for the group sample appears in Figure 4.

Follow-up t-tests were used to compare interactive effects across self-rater, peer-rater, and subordinate-rater. When rating for an administrative purpose, self-raters were significantly more lenient (M = 6.28, SD = .86) than were peer-raters (M = 5.46, SD = .82), $t(1, 60) = 7.48$, $p < .01$. Self-raters (M = 6.29, SD = .87) were also significantly more lenient than subordinate-raters (M = 5.54, SD = 1.00), $t(1, 59) = 6.98$,

$p < .01$. No significant differences were found between peer-rater and subordinate-rater under administrative conditions. In addition, no significant differences were found between rating leniency of self-rater ($M = 5.62$, $SD = .79$), peer-rater ($M = 5.71$, $SD = .73$), or subordinate-rater ($M = 5.51$, $SD = 1.20$) when rating for developmental purposes.

Results from 3x2 GLM analysis on the group member sample using discrimination between performance dimensions as the dependent variable were significant for an interaction between rater role and rating purpose $F(2, 125) = 5.60$, $p < .01$, $n^2 = .08$. As predicted, the lowest levels of discrimination between dimensions occurred for the role of self-rater when rating for an administrative purpose ($EM = .39$, $SE = .04$) as opposed to a developmental purpose ($EM = .45$, $SE = .04$). In contrast, the role of peer-rater was associated with higher levels of discrimination when rating for an administrative purpose ($EM = .57$, $SE = .04$) as opposed to developmental purpose ($EM = .46$, $SE = .03$). In a pattern similar to that of ratings of self, ratings associated with the role of subordinate-rater were less discriminant in the administrative ($EM = .55$, $SE = .04$) as opposed to developmental condition ($EM = .62$, $SE = .04$). The interaction between rater role and rating purpose on discrimination between performance dimensions for the group sample is shown in Figure 5.

Follow-up t-tests were used to compare interactive effects across self-rater, peer-rater, and subordinate-rater. When rating for an administrative purpose, self-raters were significantly less discriminant ($M = .39$, $SD = .36$) than were peer-raters ($M = .57$, $SD =$

.29), $t(1, 60) = 4.27$, $p < .01$. Self-raters ($M = .39$, $SD = .36$) were also significantly less discriminant than subordinate-raters ($M = .55$, $SD = .26$), $t(1, 59) = 3.24$, $p < .01$. No significant differences were found between peer-rater and subordinate-rater.

When rating for a developmental purpose, self-raters were found to be significantly less discriminant ($M = .45$, $SD = .32$) than subordinate-raters ($M = .62$, $SD = .34$), $t(1, 67) = 2.83$, $p < .01$. Peer-raters ($M = .46$, $SD = .27$) were also significantly less discriminant than subordinate-raters ($M = .61$, $SD = .34$), $t(1, 67) = 2.96$, $p < .01$. No significant differences were found between self-rater and peer-rater in terms of discrimination.

Main effects and Interactions: Supervisor Member Sample

A 2x2 GLM analysis was conducted on the supervisor sample ($n=50$). Rater role (self-rater, supervisor-rater) served as a within-subjects repeated factor, rating purpose (administrative, developmental) served as a between-subjects factor, and rating behaviors (rating leniency, discrimination between performance dimensions, discrimination between ratees) served as the dependent measures.

Two separate 2x2 analyses were conducted on the supervisor sample for Hypotheses one through three. The first analysis examined the effects of rater role and rating purpose on rating leniency. In the second analysis, the effects of the independent variables on discrimination between performance dimensions was examined. In addition, for Hypothesis two, discrimination between ratees served as a dependent

measure. Discrimination between ratees for the supervisor sample refers to a supervisor-rater giving ratings to subordinate ratees. Results are discussed in terms of Hypotheses one through three. A summary of results (main effects, interactions) for the supervisor sample appear in Table 7.

Main effect for rater role. Hypothesis one proposed a main effect for rater role on leniency and discrimination between performance dimensions. Means and standard deviations for rating behavior by rater role for the supervisor sample are shown in Table 1.

Results from GLM analysis on the supervisor sample found a significant effect for rater role on rating leniency, $F(1, 48) = 30.16$, $p < .01$, $n^2 = .39$. Consistent with the direction hypothesized, ratings associated with the role of self-rater ($M = 5.98$, $SD = .84$) were significantly higher than those associated with of supervisor-rater ($M = 5.40$, $SD = .90$).

GLM analysis was also conducted using discrimination between performance dimensions as the dependent variable. Contrary to the direction hypothesized, no support was found for a significant effect for rater role on discrimination between performance dimensions ($M = .41$, $SD = .29$, self-rater; $M = .49$, $SD = .24$, supervisor-rater), $F(1, 48) = 3.48$, $p > .05$.

Main effects for rating purpose. Hypothesis two proposed a main effect for rating purpose on rating leniency, on discrimination between performance dimensions, and on

discrimination between ratees. Means and standard deviations for rating purpose and rating behaviors appear in Table 8.

Through GLM analysis, a significant effect for rating purpose on leniency was found, $F(1, 48) = 12.07$, $p < .00$, $n^2 = .20$. In the direction hypothesized, higher ratings were associated with an administrative purpose for rating ($M = 6.03$, $SE = .14$) as opposed to a developmental purpose ($M = 5.37$, $SE = .13$). GLM analysis using discrimination between performance dimensions as the dependent variable failed to yield significant results, $F(1, 48) = 17$, $p > .05$.

Rater role and rating purpose interaction. Hypothesis three proposed an interaction between rater role and rating purpose on leniency and discrimination between performance dimensions.

A 2x2 GLM analysis on the supervisor sample was significant for an interaction between rater role and rating purpose on rating leniency, $F(1, 48) = 22.79$, $p < .01$, $n^2 = .32$. Findings from the interaction were based on Estimated Marginal Means (EM) that were generated for purposes of the analysis. In the direction hypothesized, the highest ratings occurred when raters rated themselves for an administrative purpose ($EM = 6.60$, $SE = .12$) as opposed to a developmental purpose ($EM = 5.42$, $SE = .12$). Ratings from supervisor-raters were similar for administrative ($EM = 5.46$, $SE = .19$) and developmental ($EM = 5.34$, $SE = .18$) purposes. The interaction between rater role and rating purpose on rating leniency for the supervisor sample is shown in Figure 6.

Follow-up t-tests were used to compare interactive effects across self-rater and supervisor-rater. When rating for an administrative purpose, self-raters were significantly more lenient (M = 6.60, SD = .50) than were supervisor-raters (M = 5.46, SD = .80), t (1,23) = 8.02, p <.01. When rating for a developmental purpose, no significant difference was found between self-rater (M = 5.42, SD = .68) and supervisor-rater (M = 5.33, SD = .99) in terms of leniency.

For the supervisor sample, a significant interaction between type of rating and purpose was also found for discrimination between performance dimensions, F(1,48) = 7.96, p < .01, η^2 = .14. As predicted, the lowest level of discrimination between dimensions occurred was associated with self-raters rating for an administrative purpose (EM = .34, SD = .06) as opposed to a developmental purpose (EM = .49, SD = .05). In contrast, ratings associated with supervisor-rater were less discriminant when given for a developmental (EM = .44, SD = .05) as opposed to an administrative purpose (EM = .55, SD = .05). The interaction between rater role and rating purpose on discrimination between performance dimensions is shown in Figure 7.

Follow-up t-tests were used to compare interactive effects across self-rater and supervisor-rater. When rating for an administrative purpose, self-raters were significantly less discriminant (M = .33, SD = .31) than were supervisor-raters (M = .55, SD = .29), t (1 ,23) = 2.80, p < .01. When rating for a developmental purpose, no significant

difference was found between self-rater (M = .49, SD = .24) and supervisor-rater (M = .44, SD = .19) in terms of discrimination.

Additional analyses. Because of the study design (not fully crossed), direct comparisons could not be made between peer and subordinate rating behavior and that of supervisor-raters. In order to make such comparisons a series of one-sample t-tests were conducted.

Supervisor ratings (M = 5.40) were found to be significantly less lenient than peer ratings (M = 5.59) ($t(1, 128) = 2.81$, $p < .01$), but not significantly different from subordinate ratings (M = 5.52) ($t(1, 127) = 1.29$, $p > .05$). In contrast, supervisor ratings were significantly less discriminant (M = .49) than subordinate ratings (M = .58) ($t(1, 127) = 3.44$, $p > .01$), but not significantly different from peer rating discrimination (M = .52) ($t(1, 128) = 1.01$, $p > .05$).

Covariates and Rating Behavior: Full Sample

GLM analysis (hypotheses four and five) and multiple regression analyses (hypotheses six through eight) were used to examine the relationship between covariates of interest and rating behavior. Findings from analyses using the full sample of participants are discussed in terms of Hypotheses four through eight. Means and standard deviations for covariate measures appear in Table 1.

Rater role, self-esteem, leniency. Hypotheses four proposed a significant interaction for rater role and self-esteem on leniency. Specifically, it was predicted that

raters in the role of self-rater and who are high in self-esteem would have more lenient ratings than raters other roles who are also high in self-esteem. For purposes of the analysis, self-esteem was grouped into "higher self-esteem" (an average score of 6 or over on the measure) and "lower self-esteem" (an average score of below 6 on the measure). A score of 6 was chosen based on descriptive analyses that determined it was the median for self-esteem scores.

Results from univariate GLM analyses using rater role (self-rater, other-rater) as a within-subjects repeated factor, self-esteem as an independent variable, and leniency as the dependent measure failed to find support for Hypothesis four, $\underline{F}$ (1, 177) = .09, $\underline{p}$ > .05.

Rater role, locus-of-control, leniency. A significant interaction for rater role and locus of control was proposed in Hypothesis five. Specifically, it was predicted that raters in the role of self-rater and who have an internal locus-of-control would give more lenient ratings than raters in other roles who also have an internal locus-of-control. For purposes of the analysis, analyses were conducted by grouping responses as either "higher locus-of-control" (an average score of 5.4 or greater) or "lower locus-of-control" (an average score below 5.4). A score of 5.4 was chosen based on descriptive analyses that determined it was the median for locus-of-control scores.

GLM analysis using rater role (self-rater, other-rater) as a within-subjects repeated factor and locus-of-control as a between-subjects factor failed to find support for Hypothesis five, $F(1, 177) = 2.61$, $p > .05$.

Covariates and self-rater role. Hypothesis six proposed that "individual characteristics" (self-esteem, locus-of-control) would account for more variance in rating behavior when the role is self-rater than would covariates representing "contextual factors" (affect, self-efficacy, rating politics, organizational commitment). Results of regression analysis are discussed in terms of leniency and discrimination between performance dimensions for self-rater.

The covariates self-esteem ($r = .20$, $p < .01$; $R^2 = .04$), self-efficacy ($r = .19$, $p < .01$; $R^2 = .04$), and organizational commitment ($r = .17$, $p < .01$; $R^2 = .03$) were each found to be significantly correlated to leniency of self-rater for the full sample. Multiple regression using stepwise selection was used to examine the relationship between covariates and leniency. Self-esteem and organizational commitment together accounted for a small amount of variance in self-rater leniency ($R^2 = .07$). Self-efficacy was removed from the analysis ($p > .05$).

Regression analysis (stepwise selection) was also conducted on self-rater discrimination between performance dimensions. Self-efficacy ($r = -.17$, $p < .05$; $R^2 = .03$) and self-esteem ($r = -.22$, $p < .01$; $R^2 = .05$) were each significantly correlated to self-rater dimension discrimination. Multiple regression using stepwise regression was used to

examine the relationship between covariates and self-rater discrimination. Only self-esteem was included in the analysis. Self-efficacy was removed from the analysis ($p > .05$).

Covariates and other-rater roles. Hypothesis seven hypothesized that individual (self-esteem, locus-of-control) and contextual(affect, self-efficacy, rating politics, organizational commitment) variables would account for significant variance in ratings associated with other-rater. Of these covariates, organizational commitment was expected to have the weakest relationship to other-rater rating behavior. Multiple regression analysis using stepwise selection was used to examine the relationship of covariates to the rating behavior of other-rater.

In support of Hypothesis 7, self-esteem ($r = .20$, $p < .01$; $R^2 = .04$), self-efficacy ($r = .23$, $p < .01$; $R^2 = .05$), and affect ($r = .36$, $p < .01$; $R^2 = .13$) each were related to other-rater leniency. Multiple regression analysis using stepwise selection was used to examine the relationship between the covariates and rating behavior of other-rater. Self-efficacy and affect together accounted for a moderate amount of variance in other-rater leniency ($R^2 = .17$). Self-esteem was removed from analysis ($p > .05$). No covariates were significantly related to discrimination between performance dimensions for other-rater. A summary of Regression results for self and other-rater appears in Table 9.

Covariate influence on the rater role, rating purpose, rating behavior relationship. In Hypothesis eight, it was predicted that when the effects of covariates were controlled

for, the relationship between rater role, rating purpose, and rating behavior would become less significant. Findings are discussed in terms of the full sample. Refer to Table 10 for a summary of results from GLM analysis controlling for covariates.

A 2x2 GLM analysis was conducted controlling for the effects of self-esteem, self-efficacy, affect, and organizational commitment. Covariates controlled for were those that were significantly related to the rating behavior of self-rater and other-rater in the full sample. In the analysis, rater role (self-rater, other-rater) served as a within-subjects repeated factor and rating purpose (administrative, developmental) served as the between-subjects factor. The analyses were conducted twice, once using rating leniency as the dependent measure and once using discrimination between performance dimensions. Results are discussed in terms of each type of rating behavior.

The 2x2 analysis described above was conducted using leniency as the dependent measure. When the effects of covariates were controlled for, the relationship between rater role and rating leniency for the full sample went from significant ($F(1,175) = 61.22$, $p < .01$, $n^2 = .26$) to non- significant ($F(1, 175) = 1.55$, $p > .05$. Controlling for covariates did not change the main effect for rating purpose on leniency, which remained significant. In addition, controlling for covariates did not change the interaction between rater role, rating purpose, and leniency. The interaction remained significant.

A 2x2 GLM analysis was also conducted using discrimination between performance dimensions as the dependent measure. When the effects of covariates were

controlled for, the relationship between rater role and discrimination between performance dimensions for the full sample went from significant ($F(1, 175) = 14.43$, $p < .01$, $n^2 = .08$) to non- significant ($F(1, 172) = .49$, $p > .05$). Controlling for covariates did not change the relationship between rating purpose and dimension discrimination (remained non-significant), or the interaction between rater role and rating purpose on dimension discrimination (remained significant).

Covariates and Rating Behavior: Group Member Sample

GLM analysis (hypotheses four, five, and eight) and multiple regression analyses (hypotheses six and seven) were used to examine the relationship between covariates of interest and rating behavior. Findings from the analyses using the group sample of participants (those who served in the position of group members in the experiment) are discussed in terms of Hypotheses four through eight.

Rater role, self-esteem, and leniency. In Hypothesis one, a significant interaction for rater role and self-esteem on leniency was proposed. Results from GLM analysis using rater role (self-rater, peer-rater, subordinate-rater) as a within-subjects repeated factor, self-esteem (grouped as high or low) as an independent variable, and leniency as the dependent measure failed to find support for Hypothesis four for the group member sample, $F(2 , 125) = .30$, $p > .05$.

Rater role, locus-of-control, and leniency. A significant interaction for rater role and locus-of-control was proposed in Hypothesis 5. Results from GLM analysis using

rater role (self-rater, peer-rater, subordinate-rater) as a within-subjects repeated factor, internal locus-of-control (grouped as high or low) as an independent variable, and leniency as the dependent measure failed to find support for Hypothesis five for the group member sample, $F(2, 125) = .74$, $p > .05$.

Covariates and self-rater role. Hypothesis six proposed that "individual characteristics" (self-esteem, locus-of-control) would account for more variance in rating behavior when the role is self-rater than would covariates representing "contextual factors" (affect, self-efficacy, rating politics, organizational commitment). Results of multiple regression analysis are discussed in terms of leniency and discrimination between performance dimensions for self-rater.

Multiple regression analysis using stepwise selection was conducted using leniency of self-raters from the group member sample as the dependent measure. The covariates self-esteem ($r = .25$, $p < .01$; $R^2 = .06$), self-efficacy ($r = .18$, $p < .05$; $R^2 = .03$) and organizational commitment ($r = .18$, $p < .05$; $R^2 = .03$) each were found to be significantly related to leniency of self-raters in the group member sample. A moderate amount of total variance ($R^2 = .09$) in leniency of self-rater was accounted for by self-esteem and organizational commitment. Self-efficacy was removed from analysis ($p > .05$).

Self-esteem ($R^2 = .07$) and self-efficacy ($R^2 = .03$) each accounted for a small amount of variance in the discrimination between performance dimensions associated

with self-rater for the group member sample. Only self-esteem was included in the regression analysis. Self-efficacy was removed from analysis ($p < .05$).

Covariates and other-rater roles. In Hypothesis seven, it was predicted that individual (self-esteem, locus-of-control) and contextual(affect, self-efficacy, rating politics, organizational commitment) variables would account for significant variance in ratings associated with other-rater (peer-rater, subordinate-rater). Of these covariates, organizational commitment was expected to have the weakest relationship to other-rater rating behavior.

The influence of covariates on leniency associated with peer-rater and subordinate-rater roles were examined through multiple regression analysis using stepwise selection. Affect ($r = .24$, $p < .01$; $R^2 = .06$), self-efficacy ($r = .25$, $p < .01$; $R^2 = .06$), and self-esteem($r = .29$, $p < .01$; $R^2 = .09$), were each found to account for small to moderate amounts of variance in leniency associated with peer-rater. Affect and self-esteem together were found to account for a moderate amount of variance in peer-rater leniency ($R^2 = .12$). Self-efficacy was removed from the analysis ($p > .05$).

Self-esteem ($r = -.26$, $p < .01$; $R^2 = .07$) was the only covariate to account for variance in discrimination between performance dimensions for the peer-rater role.

Self-esteem ($r = .20$, $p < .01$; $R^2 = .04$), politics ($r = -.18$, $p < .05$; $R^2 = .03$), affect ($r = .24$, $p < .01$; $R^2 = .06$), and efficacy ($r = .22$, $p < .01$; $R^2 = .05$) each accounted for small amounts of variance in leniency associated with subordinate-rater. Affect, self-

efficacy, and politics together accounted for a moderate amount of variance in leniency of subordinate-rater ($R^2 = .14$). Self-esteem was removed from the analysis ($p > .05$). No covariates were significantly related to dimension discrimination for subordinate-raters.

In the direction hypothesized, organizational commitment did not account for a significant amount of variance in any rating behavior for peer or subordinate-rater. A summary of Regression results for self, peer, and subordinate-raters appears in Table 11.

Covariate influence on the relationship between rater role, rating purpose and rating behavior. It was hypothesized that when the effects of covariates were controlled for that the relationship between rater role (self-rater, peer-rater, subordinate-rater), rating purpose (administrative, developmental), and rating behavior would be less significant. The hypothesis was tested using the group member sample. A summary of results for the group member sample appears in Table 12.

A 3x2 GLM analysis was conducted controlling for the effects of self-esteem, self-efficacy, affect, and politics. Covariates controlled for were those that were significantly related to the rating behavior of self-rater and other-rater (peer-rater, subordinate-rater) in the full sample. In the analysis, rater role (self-rater, peer-rater, subordinate-rater) served as a within-subjects repeated factor and rating purpose (administrative, developmental) served as the between-subjects factor. The analyses were conducted twice, once using rating leniency as the dependent measure and once

using discrimination between performance dimensions. Results are discussed in terms of each type of rating behavior.

The 3x2 analysis described above was conducted using leniency as the dependent measure. When the effects of covariates were controlled for, the relationship between rater role and rating leniency for the group sample went from significant ($F(2, 125) = 17.03$, $p < .01$) to non- significant ($F(2, 121) = .37$, $p > .05$). Controlling for covariates did not change the main effect for rating purpose on leniency. The main effect for purpose remained nonsignificant. In addition, controlling for covariates did not change the interaction between rater role, rating purpose, and leniency. The interaction remained significant.

A 3x2 GLM analysis was also conducted using discrimination between performance dimensions as the dependent measure. When the effects of covariates were controlled for, the relationship between rater role and discrimination between performance dimensions for the group sample went from significant ($F(2, 125) = 9.15$, $p < .01$) to non- significant ($F(2, 121) = 1.68$, $p > .05$). The main effect for purpose remained nonsignificant. The interaction between rater role, rating purpose, and leniency remained significant after controlling for covariates.

Covariates and Rating Behavior: Supervisor Member Sample

GLM analysis (hypotheses four, five, and eight), and hierarchical regression analyses (hypotheses six and seven) were used to examine the relationship between

covariates of interest and rating behavior. Findings from the analyses using the supervisor sample of participants (those who served in the position of supervisor member in the experiment) are discussed in terms of Hypotheses four through eight.

Rater role, self-esteem, and leniency. Hypothesis one proposed a significant interaction for rater role and self-esteem on leniency. Results from GLM analysis using rater role (self-rater, supervisor-rater) as a within-subjects repeated factor, self-esteem (grouped as high and low) as an independent variable, and leniency as the dependent measure failed to find support for Hypothesis four for the supervisor member sample, $F(1, 48) = .74$, $p > .05$.

Rater role, locus-of-control, and leniency. A significant interaction for rater role and locus-of-control was proposed in Hypothesis five. Results from GLM analysis using rater role (self-rater, supervisor-rater) as a within-subjects repeated factor, internal locus-of-control (grouped as high and low) as an independent variable, and leniency as the dependent measure failed to find support for Hypothesis five for the supervisor member sample, $F(1, 48) = 1.15$, $p > .05$.

Covariates and self-rater role. Hypothesis six proposed that those covariates representing "individual characteristics" would account for more variance in rating behavior when the role is self-rater than would those representing "contextual factors". Results of regression analysis for the supervisor sample may be discussed in terms of leniency and discrimination between performance dimensions for self-rater.

Multiple regression analysis using stepwise selection was conducted using leniency of self-raters from the supervisor sample as the dependent measure. In the direction hypothesized, only locus-of-control ($r = -.30$, $p < .01$) accounted for significant variance in leniency ($R^2 = .08$) of self-rater from the supervisor member sample.

Multiple regression analysis using stepwise selection were also conducted using discrimination between performance dimensions associated with self-raters from the supervisor sample as the dependent measure. Locus-of-control ($r = .30$, $R^2 = .09$) and affect ($r = -.30$, $R^2 = .09$) were significantly related to dimension discrimination. When entered into regression analysis, locus-of-control and affect together accounted for a moderate amount of variance ($R^2 = .19$) in dimension discrimination.

Covariates and other-rater roles. Hypothesis seven hypothesized that individual and contextual variables would account for significant variance in ratings associated with other-rater (supervisor-rater). Of these covariates, organizational commitment was expected to have the weakest relationship to other-rater rating behavior.

The relationships between covariates and rating behavior were examined for supervisor-rater role through simple regression analysis. Only affect ($r = .53$, $p < .01$) was found to be significantly related to leniency associated with supervisor-rater in the supervisor member sample. A large amount of variance in supervisor-rater leniency was accounted for by affect ($R^2 = .29$). No covariates were significantly related to

discrimination between performance dimensions for supervisor-rater. A summary of Regression results for both self and supervisor-raters appears in Table 13.

Covariate influence on the relationships between rater role, rating purpose, and rating behavior. It was hypothesized that when the effects of covariates were controlled for that the relationship between rater role, rating purpose, and rating behavior would be less significant. This hypothesis was tested using the supervisor member sample. A summary of results for the supervisor sample appears in Table 14.

A 2x2 GLM analysis was conducted controlling for the effects of affect and locus-of-control. Covariates controlled for were those that were significantly related to the rating behavior of self-rater and other-rater (supervisor-rater) in the supervisor sample. In the analysis, rater role (self-rater, supervisor-rater) served as a within-subjects repeated factor and rating purpose (administrative, developmental) served as the between-subjects factor. The analyses were conducted twice, once using rating leniency as the dependent measure and once using discrimination between performance dimensions. Results are discussed in terms of each type of rating behavior.

The 3x2 analysis described above was conducted using leniency as the dependent measure. When the effects of covariates were controlled for, the relationship between rater role and rating leniency for the supervisor sample went from significant (F (1,48) = 30.16, $p < .01$) to non- significant ($F(1, 46) = 1.81$, $p > .05$). Controlling for covariates did not change the main effect for rating purpose on leniency. The main effect for

purpose remained significant. In addition, controlling for covariates did not change the interaction between rater role, rating purpose, and leniency. The interaction remained significant.

A 2x2 GLM analysis was also conducted using discrimination between performance dimensions as the dependent measure for the supervisor sample. When covariates were controlled for, no changes occurred in the relationship between rating purpose and dimension discrimination (remained nonsignificant) or for the interaction of rating purpose, rater role, and dimension discrimination (remained significant).

CHAPTER IV

DISCUSSION

The current study had two primary purposes. First, the current study sought to examine the rating behaviors associated with self-rater and other-rater (peer-rater, subordinate-rater and supervisor-rater) roles. The focus of the study was to examine differences in ratings given as a self-rater with those given to others. In this respect, the current study provided a unique look at self and other rating discrepancy not previously examined in past literature. Secondly, the current study sought to examine a variety of individual and contextual variables for their relationship to rating behavior associated with the role of self-rater and other-rater (peer-rater, subordinate-rater, supervisor-rater). It was proposed that individual characteristics would be especially important to the rating behaviors of self-rater. Findings from the current study are discussed in terms of contributions to literature and organizational implications.

Main Effects and Interactions

Hypotheses one through three examined the relationship between rater role, rating purpose, and rating behavior. Findings are discussed in terms of the full sample (ratings from all participants, grouped as ratings from self-rater or other-rater), the group sample (participants who served in experimental task as a group member and served in the roles of self-rater, peer-rater, and subordinate-rater), and the supervisor sample (participants

who served in experimental task as a supervisor and served in the roles of self-rater and supervisor-rater).

Main effect for rater role. Hypothesis one predicted a main effect for rater role on rating behavior. Specifically, it was predicted that ratings from the role of self-rater would be more lenient and less discriminant that ratings associated with any other-rater role (peer, subordinate, and supervisor). Findings from the full, group, and supervisor samples supported Hypothesis one.

The role of self-rater was associated with the most lenient ratings of all rater roles and across all samples (full, group member, and supervisor member). In the full sample, ratings associated with the self-rater were found to be significantly more lenient (M= 5.95) than those associated with the role of other-rater (M=5.51). Similarly, ratings associated with the role of self-rater in the group member sample were found to be significantly more lenient (M=5.93) than either ratings associated with peer-raters (M=5.60) or those associated with subordinate-raters (M=5.53). Ratings associated with the role of self-rater in the supervisor member sample were also found to be significantly more lenient (M=5.98) than ratings associated with the supervisor-rater role (M=5.40).

Support was also found for Hypothesis one when discrimination between performance dimensions was examined. As predicted, ratings from self-rater in the full sample were found to be significantly less discriminant between performance dimensions (M=.42) than were ratings from other-rater (M=.53). Likewise, self-ratings in the group

member sample were significantly less discriminant between performance dimensions (M=.42) than either ratings associated with peer-raters (M=.51) or those associated with subordinate-raters (M=.58). Interestingly, ratings from the peer-rater role were found to be significantly less discriminant than those associated with the subordinate-rater role. Contrary to what was hypothesized, no significant effect for rater role on discrimination between performance dimensions was found in the supervisor member sample.

When findings from all three samples are considered, consistent support was found for the ratings from self-raters being more lenient and less discriminant than those from any other rating source. These findings are unique in that they examine a comparison that has not been systematically examined in past research, that of self-ratings with ratings given to others. Previous research has supported that self-ratings tend to be discrepant from those received from others (e.g., Harris & Schaubroeck, 1988). Findings from the current study confirm that the same pattern of results may be found when comparing self-ratings with ratings given to others.

Findings from the current study support that some differences may be expected in terms of rating behavior when comparing peer, subordinate, and supervisor-raters. Supervisor ratings were significantly less lenient than peer ratings, but similar to those of subordinates. In terms of dimension discrimination, both supervisor and peer ratings were significantly less discriminant than subordinate ratings. While these findings indicate that differences between raters other than self may be expected, in general they are not as

dramatic as those between self-rater and other-rater. This pattern of results is similar to that found in previous literature that supports significant differences primarily between self and other-raters, but not between other-raters themselves (e.g. Harris & Schaubroeck, 1988; Mabe & West, 1982).

Interestingly, the largest discrepancy between self and other rating leniency existed in the supervisor sample between self-ratings and supervisor ratings. While self ratings in the supervisor sample were similar to self-ratings from the group sample, supervisor ratings (supervisor rating a subordinate) were less lenient than were peer ratings or subordinate ratings (subordinate rating a supervisor). While this finding is unique in comparing ratings given to self with those given to other, it is supported by previous research. Thornton (1968) found discrepancies between self and supervisor ratings. When comparing self, peer, and supervisor ratings, Harris and Schaubroeck (1988), found the largest discrepancies between ratings were between self and supervisor. Harris and Schaubroeck (1988) did not examine subordinate ratings and the current study contributes to the literature in its inclusion of this rater role.

Across samples, a stronger effect was found for rater role on leniency ($\underline{n}^2$=.26, full sample; $\underline{n}^2$=.21, group sample; $\underline{n}^2$=.39, supervisor sample) than was found for dimension discrimination ($\underline{n}^2$=.08, full sample; $\underline{n}^2$=.13, group sample; $\underline{n}^2$= n.s.). It may be that the role a rater is in may affect his or her ability to distinguish between performance dimensions in ratings to a lesser extent than it effects the degree to which he or she is

lenient in rating. Previous researchers (e.g., Landy & Farr, 1983) have found leniency to be the most commonly encountered rating error.

Main effect for rating purpose. Hypothesis two predicted a main effect for rating purpose on rating behavior. Specifically, it was predicted that the highest and least discriminant (between ratees and between performance dimensions) ratings would be found in the administrative condition as opposed to developmental condition for all rater roles. While support was found for a main effect for rating purpose on leniency, no support was found for a main effect for purpose on either discrimination between performance dimensions or between ratees. Findings may be discussed in more detail in terms of the full, group, and supervisor sample.

In support of Hypothesis two, findings from the full sample of participants found ratings associated with the administrative condition to be significantly more lenient (M= 5.96) than those associated with the developmental condition (M=5.50). Similarly, findings from the supervisor sample of participants showed ratings associated with the administrative condition to be significantly more lenient (M= 6.03) than those associated with the developmental condition (M=5.37). Interestingly, no significant effect for rating purpose on rating leniency was found for the group sample. In addition, no significant purpose effect was found for discrimination between performance dimensions or discrimination between ratees for any sample.

Findings from the full and supervisor samples supported an effect for purpose on leniency. The relationship was in the direction supported by previous research that has found more lenient and less discriminant ratings to be associated with administrative purposes (e.g., Zedeck & Cascio, 1982).

One implication of findings from the current study is that depending on whether the rater serves in a supervisory (supervisor sample) or nonsupervisory position (group sample), a difference in effect that purpose has on leniency may be expected. When rating leniency from group members were considered, purpose had no effect. In contrast, when rating leniency from the supervisor sample of participants was considered, purpose had a large effect ($\underline{n}^2$=.20).

Interaction of rater role and rating purpose. Hypothesis three predicted an interaction for rater role and rating purpose on rating behavior. Specifically, it was predicted that more lenient and less discriminant (between performance dimensions) ratings would result from ratings from the self-rater role done for administrative purposes than ratings from the self-rater role done for developmental purposes. In support of Hypothesis three, ratings from self-raters done for administrative purposes were found to be more lenient and less discriminant than those associated with developmental purposes for each sample. Findings from the full, group, and supervisor samples supported hypothesis three.

In support of Hypothesis three, findings from each sample found ratings associated with self-rater in the administrative condition to be significantly more lenient (EM= 6.44, full sample; EM=6.29, group sample; EM=6.60, supervisor sample) than those associated with self-rater in the developmental condition (EM=5.52, full sample; EM=5.62, group sample; EM=5.42, supervisor sample).

Findings from all samples supported Hypothesis three in terms of discrimination between performance dimensions. Self-ratings done for an administrative purpose were found to be less discriminant (EM=.36, full sample; EM= .39, group sample; EM= .34, supervisor sample) than those associated with a developmental purpose (EM= .47, full sample; EM=.45, group sample; EM= .49, supervisor sample). Although no specific hypotheses were made regarding rater roles other than self, findings may be discussed in terms of these other roles.

No significant differences were found between ratings done for administrative purposes and those done for developmental purposes in terms of either leniency for other-rater (full sample) peer-rater, subordinate-rater, or supervisor-rater. It is of interest to note however, that while not significantly different, peer-ratings for a developmental purpose were more lenient than those done for an administrative purpose. This finding is in a direction opposite of previous research (e.g. Farh et al, 1991) that found peer ratings for administrative purposes to be more lenient than those for developmental purposes. It may be that in the current study peers gave less lenient ratings to coworkers in the

administrative condition because they were competing against them for a monetary reward.

In terms of discrimination, a significant difference was found only for peer-raters. In contrast to findings involving self-raters, peer-raters were significantly more discriminant when rating for an administrative (M=.57) as opposed to developmental purpose (M=.46). Previous research (e.g., Farh et al., 1991) has found the opposite effect, administrative ratings being associated with less discrimination than developmental ratings. As mentioned above, competition for rewards may have contributed to findings in the current study. Ratings may also be discussed in terms of comparisons by purpose and across rater roles for each sample.

In the full sample, self-ratings for an administrative purpose were found to be significantly different from those of other-rater in terms of leniency (M=6.37, self; M=5.48, other-rater) and dimension discrimination (M=.37, self; M=.56, other-rater). When rating for a developmental purpose, no significant differences between self and other-rater were found for either leniency or discrimination.

In the group sample, self-ratings for an administrative purpose were found to be significantly different from those of both peer and subordinate-raters in terms of leniency (M=6.28, self; M=5.46, peer-rater; M=5.54, subordinate-rater) and dimension discrimination (M=.39, self; M=.57, peer-rater; M=.55, subordinate-rater). No significant

differences existed between peer-rater and subordinate-rater administrative ratings in terms of leniency or discrimination.

When rating for a developmental purpose, both self-raters (M=.45) and peer-raters (M=.46) were found to be significantly less discriminant than subordinate-raters (M=.62). No significant differences between self, peer, and subordinate rating leniency was found for the developmental condition.

In the full sample, self-ratings for an administrative purpose were found to be significantly different from those of supervisor-rater in terms of leniency (M=6.60, self; M=5.46, supervisor-rater) and dimension discrimination (M=.33, self; M=.55, supervisor-rater). When rating for a developmental purpose, no significant differences between self and supervisor-rater were found for either leniency or discrimination.

In support of Hypothesis three and consistent with previous literature, self-ratings associated with an administrative purpose were significantly more lenient and less discriminant than those associated with a developmental purpose across samples (e.g., Campbell & Lee, 1988). In addition, self-ratings done for an administrative purpose were significantly more lenient and less discriminant than those from any other-rater role also done for administrative purposes. Previous researchers (e.g., Longenecker et al., 1987) have proposed that self-raters are driven to give themselves lenient ratings under administrative conditions due to a desire to protect themselves from negative consequences associated with low ratings.

Interestingly, when rating for developmental purposes no significant differences were found between self-ratings and those from other-rater roles in terms of leniency. In terms of discrimination related to a developmental purpose, only one significant difference was found. Subordinate ratings were significantly more discriminant under developmental conditions than those from self or peer-raters. In general, findings from the current study suggest that rating purpose is of particular importance to rating behavior when the rater role is self.

Implications for the organization. Findings from the current study support that the types of ratings (in terms of leniency and discrimination) coming from an organization's evaluation process may be expected to be influenced by both who is doing the rating (rater role) and the reason behind the rating (rating purpose). Implications for the organization may be discussed in terms of design and implementation of performance evaluation systems.

Organizations using multi-rater processes may want to carefully consider the importance given to ratings associated with different rater roles will have in contributing to decisions made about the ratee. Of special concern may be ratings from self-raters done for administrative purposes. In the current study, self-ratings were found to be especially discrepant from those associated with other-rater roles in terms of leniency and discrimination. This effect was especially salient when rating for administrative purposes. Interestingly, when rating for developmental purposes, self-ratings shared

similar levels of leniency and discrimination with those associated with those associated with peer, supervisor, and subordinate-raters.

These findings suggest that organizations may want to be especially wary of placing much consideration on self-ratings done for an administrative purpose, but may want to give them the same consideration as those from other-raters when rating is for developmental purposes. The true value of self-ratings done for administrative purposes may be in their ability to allow individuals to contribute to their evaluations, but not as determinants in any administrative decisions. When ratings are being used for administrative purposes, those from sources other than self may be viewed as superior to self-ratings.

Findings regarding "other-raters" (full sample), found ratings from "others" done for administrative purposes to be less lenient and more discriminant than those from self-raters done for administrative purposes. While peer, supervisor, and subordinate-raters were more similar than different, a few significant differences did exist. While peer, subordinate, and supervisor-raters were not found to differ in terms of leniency, ratings from subordinate-rater (subordinate rating a supervisor) were the most discriminant. In addition, peers were less discriminant when rating for an administrative purpose then for a developmental purpose. As some differences in rating behaviors of those in rater roles other than self may be expected, organizations may want to take care not to rely on the ratings associated with any single "other-rater" role. An approach that looks for

convergence among the "other-rater" roles or that averages ratings across "other-raters" may be superior in value to one that relies on ratings from any one source.

It may be argued that other-raters (as a group) may be expected to provide more accurate ratings than those provided by self. While some might find this reason to abandon self-ratings, their importance to the evaluation process in allowing the individual to have input into their evaluation is critical and should not be eliminated. In addition, when done for developmental purposes self-ratings were found in the current study to be quite similar to ratings associated with other-rater roles.

Organizations may want to develop evaluation systems with purpose in mind. Findings from the current study may be especially important given a growing trend in organizations to use ratings from sources other than supervisor for administrative decisions (i.e. pay and promotion; London & Smithers, 1995). When organizations are using ratings for an administrative purpose they may want to do so with the understanding that in general, such ratings may be more lenient than had they been collected solely for developmental purposes.

Three general guidelines have been proposed for organizations using 360 degree feedback processes for multiple purposes: 1) Base developmental decisions on feedback regarding competence and behaviors, 2) base administrative decisions on feedback related to accomplishments and results, 3) protect confidentiality by limiting the portion of feedback that supervisors can see (Waldman, Atwater, & Antonioni, 1998).

Organizations that follow such guidelines and who look for convergence among ratings from multiple sources may be expected to have better outcomes from their evaluation processes than organizations that do not.

Covariates and Rating Behavior

Hypotheses four through eight addressed the relationship between rating behavior and a number of individual (self-esteem, locus-of-control) and contextual (self-efficacy, affect, politics, organizational commitment) variables. Each hypothesis may be discussed in terms of findings from the full, group samples, and supervisor member samples.

Rater role, self-esteem, locus-of-control and leniency. Hypothesis four proposed a significant interaction between rater role and self-esteem on leniency. Specifically, it was predicted that raters in the role of self-rater and who are high in self-esteem would have more lenient ratings than raters in the role of other-rater who are also high in self-esteem.

Hypothesis five proposed a significant interaction between rater role and locus-of-control on leniency. Specifically, it was predicted that raters in the role of self-rater and who have an internal locus-of-control would give more lenient ratings than raters in the role of other-rater who also have an internal locus-of-control.

Contrary to what was predicted, findings using the full, group member and supervisor member samples were not significant for an interaction for either Hypothesis four or Hypothesis five. The relationships between rater role, self-esteem and leniency

and between rater role, locus-of-control and leniency were similar for self, peer, subordinate, and supervisor raters.

Covariates and self-rater role. Hypothesis six proposed that "individual characteristics" (self-esteem, locus-of-control) would account for more variance in rating behavior (leniency, discrimination between performance dimensions) for the role of self-rater than would "contextual factors"(affect, self-efficacy, rating politics, and organizational commitment). Contrary what was hypothesized, similar amounts of variance in ratings from self-raters were accounted for by both individual and contextual factors. One exception to this was found in the supervisor member sample. Results may be discussed in terms of each covariate and variance accounted for in self ratings.

Self-esteem was found to account for significant variance in leniency in self-raters in both the full ($\underline{R}^2$=.04) and group samples ($\underline{R}^2$=.06). Findings are consistent with research that has found individuals high in self-esteem to give themselves more lenient self-ratings (Farh & Dobbins, 1989). Self-esteem also accounted for variance in discrimination between performance dimensions for the self-raters in the full sample ($\underline{R}^2$=.05) and the group sample ($\underline{R}^2$=.07). In both instances the relationship was negative such that an increase in self-esteem was related to a decrease in discrimination.

Locus-of-control accounted for a significant amount of variance in leniency (R^2=.08) and discrimination (R^2=.09) associated with self-rater in the supervisor sample. The relationship between leniency and locus-of-control was negative (a more internal

locus-of-control being related to less lenient ratings), while the relationship between discrimination and locus-of-control was positive (a more internal locus-of-control being related to more discriminant ratings). Findings regarding leniency differ from those of previous research (e.g., Levy, 1991) that found individuals with internal locus-of-control to report higher levels of performance than those with external locus-of-control. As locus-of-control was not significantly related to self-ratings from the group sample, it may be that supervisory position contributed to the relationship between locus-of-control and self-rater leniency in the current study.

Self-efficacy accounted for variance in self-rater leniency for both the full (R^2=.04) and the group (R^2=.03) samples. In addition, self-efficacy accounted for variance in discrimination between performance dimensions for both the full (R^2=.03) and the group (R^2=.03) samples. The relationship between self-efficacy and dimension discrimination for self-raters in both the full and group samples was negative. An increase in self-efficacy was related to a decrease in discrimination. These findings are supported by theory (Murphy & Cleveland, 1991) that suggests that due to their frequent access to performance related behaviors and results of behaviors, high self-efficacy in self-raters may be related to more lenient and less discriminant ratings.

Affect accounted for variance in dimension discrimination associated with self-raters from the supervisor member sample (R^2=.09). The relationship was negative such that greater liking for the ratee was associated with less discrimination. This relationship

is consistent with previous research that has found liking of a ratee to be related to leniency, halo, and low interrater agreement (Kingstrom & Mainstone, 1983).

Organizational commitment accounted for similar amounts of variance in self-rater leniency in both the full (R^2=.03) and group (R^2=.03) samples as was accounted for by individual characteristics (i.e., self esteem). Greater commitment to the organization was related to more lenient ratings. Theory on organizational commitment (e.g., Mowday, Porter, & Steers, 1982) would predict that greater commitment to the organization would be related to less leniency in performance ratings. Previous research had not examined the influence of organizational commitment on self ratings.

Contrary to what was hypothesized, findings support that the level of variance accounted for in self-rating behavior by individual and contextual factors was comparable. Interestingly, the types of variables of significant to self-rating behavior differed between samples. Self-esteem, self-efficacy and organizational commitment were significantly related to rating behavior of full and group self-raters. In support of Hypothesis six, only locus-of-control accounted for significant variance in self-rating behavior from the supervisor sample. It may be that the importance of individual characteristics to rating behavior is in part determined by the organizational level of the self-rater (supervisory or nonsupervisor). Of all the covariates of interest, only politics was not significantly related to any self-rater rating behavior. It may be that when rating

one's own performance, the degree to which others are perceived as manipulating ratings is not as important as it might be when rating the performance of others.

Covariates and other-rater roles. Hypothesis seven proposed that individual and contextual variables would account for significant variance in the rating behavior of raters other than self (peer-rater, subordinate-rater, supervisor-rater). In addition, it was hypothesized that of all covariates, organizational commitment would account for the least amount of variance in other-rating behavior of all covariates.

In support of Hypothesis seven, all but one of the hypothesized covariates (locus-of-control) was found to account for variance for at least one rating behavior (leniency, discrimination) and for at least one rater role (peer, subordinate, supervisor). Providing additional support for Hypothesis seven. organizational commitment did not account for any variance in the rating behavior of any other-rater role. Results may be discussed in terms of each covariate and variance accounted for.

Self-esteem was found to account for a variance in leniency associated with the role of other-rater (full sample) (R^2=.04), and with the roles of peer-rater (R^2=.09) and subordinate-rater role (R^2=.04) (group sample). Higher self-esteem was related to increased leniency. In addition, self-esteem was related to discrimination for peer-raters (R^2=.07). The relationship was negative such that an increase in self-esteem was related to a decrease in discrimination. As was discussed in the preceding section, these findings were consistent with those found in previous research examining self-raters. However,

previous research has not examined the relationship between self-esteem and raters other than self. It may be that self-esteem is not only influential in how one rates his or her own performance, but the performance of others as well.

Self-efficacy accounted for variance in leniency associated with the role of other-rater (R^2=.05)(full sample), and the roles of peer-rater (R^2=.06) and subordinate-rater (R^2=.05) (group sample). Higher levels of self-efficacy were related to an increase in leniency. As was discussed in the preceding section, past research supports that greater self-efficacy may be related to increased leniency for self-raters. However, literature suggests that the opposite may be true for those in other-rater roles. Based on theory involving self-efficacy, other-raters who have a good understanding of the appraisal process may be expected to be less lenient in their ratings (Gardner & Pierce, 1998). It is interesting that self-efficacy was related to leniency for other-raters from the group sample (peer and subordinate), but not to supervisor leniency in the supervisor sample. It may be that that a unique element of organizational level (supervisory or nonsupervisory position) contributes to the influence that self-efficacy has on rater leniency.

Affect accounted for significant variance in leniency associated with the role of other-rater (R^2=.13; full sample), and with the roles of peer-rater (R^2=.06) and subordinate-rater (R^2=.06; group sample). In addition, it was the only covariate significantly related to leniency in the role of supervisor-rater (R^2=.29). Interestingly, the amount of variance accounted for by affect in supervisor leniency was the largest amount

accounted for by any single covariates across rater roles and rating behaviors. In a previous study of supervisor-raters, higher levels of liking for a ratee were found to be related to higher ratings (Kingstrom & Mainstone, 1985).

Politics was found to account for variance for only one rating behavior and one rating role, that of leniency (R^2=.03) associated with the role of subordinate-rater. Interestingly, this relationship was negative (r = -.18) such that a belief that raters manipulate ratings was related to lower ratings. Previous research by Cleveland et al. (2000) has demonstrated the opposite effect (a belief that other-raters do manipulate ratings related to more lenient ratings) using a sample of primarily supervisory ratings (supervisor rating a subordinate). It may be that when subordinate-raters believe there is an absence of rating manipulation (perhaps among their supervisors), they may be inclined to rate their supervisor more favorably. While previous research has not addressed this relationship directly, previous research has supported that raters who perceive their organization's culture as less political are more inclined to provide accurate performance evaluations (Gioia & Longenecker, 1994).

In support of Hypothesis seven, organizational commitment did not account for significant variance in the rating behavior of any other-rater (peer-rater, subordinate-rater, supervisor-rater). These findings are supported by previous research (e.g., Cleveland et al., 2000) that found factors which are more removed from raters ("distal" variables) to

account for less variance in rating behavior than those that more directly influence raters ("proximal" variables).

As was found with self-rating behavior, both individual and contextual factors accounted for significant variance in the rating behavior of other-raters (other-rater, peer-rater, subordinate-rater, supervisor-rater). Interestingly, depending on the rater role (other-rater, peer-rater, subordinate-rater, supervisor-rater) different covariates were found to have importance to rating behavior.

Self-esteem, affect, and self-efficacy were all related to both peer and subordinate rating leniency. In addition, rating politics was related to subordinate-rater leniency. For supervisor-raters, only affect was related to rating leniency. For other-rater (full sample), both self-esteem and self-efficacy were related to leniency. While organizational commitment was found to be significantly related to self-rater leniency, it was not related to the rating behavior of any other-rater roles. It may be that the rating behavior of other-raters will be influenced by individual and contextual variables to differing extents based on the role the other-rater is in, much as covariate influence on self-raters varied depended on supervisor or nonsupervisor status.

One interesting finding in the current study is that some variables were significantly related to rating behavior for both self and other-rater roles, while others were not. Self-esteem, self-efficacy, and affect were related to rating behavior for both self and other-rater roles. Locus-of-control and organizational commitment were related

only to self-raters. Politics was related only to a other-rater role (subordinate-rater). It may be that some covariates will have a general importance regardless of rater role, while the significance of others will be dependent on the role of the rater.

Covariate influence on the relationships between rater role, rating purpose, and rating behavior. Hypothesis eight proposed that when the effects of covariates were controlled for that the relationship between rater role (self-rater, peer-rater, subordinate-rater, supervisor-rater), rating purpose (administrative, developmental), and rating behavior (leniency, discrimination between performance dimensions, discrimination between ratees) would be less significant. Findings provide mixed support for Hypothesis eight.

In support of Hypothesis eight, when the effects of covariates were controlled for in the full sample, group, and supervisor samples, the relationship between rater role (self-rater, other-rater) and rating leniency went from significant to non-significant in each case. In addition, controlling for covariates in each of the three samples resulted in the relationship between rater role and discrimination between performance dimensions to go from significant to nonsignificant. Contrary to what was hypothesized, controlling for covariates did not change the relationship between rating purpose and leniency or discrimination, nor the interaction between role, purpose and leniency or discrimination for the full, group, or supervisor samples. Covariates may be especially important for

understanding the relationship between rater role and rating behavior, but not for rating purpose and rating behavior.

Organizational implications. Findings regarding covariates suggest that organizations need to be aware of the potential influence that both individual and contextual variables may have on performance ratings. The greater an organization's understanding of factors that may contribute to rating leniency and discrimination, the better able they may be to identify and address such issues. Feelings about one's worth (self-esteem), the level to which one feels knowledgeable about the rating process (self-efficacy), and the degree to which a rater likes a ratee (affect) were found to be significantly related to the rating behavior of both self and other-raters. Each of these covariates are discussed in terms of implications for rater training.

While organizations may be limited in the degree to which they may dictate how worthy an individual feels or how well an individual likes a ratee, they can develop an awareness in individuals as to the potential effect such beliefs may have on ratings they give themselves and others. Rater training programs may be used to inform raters about the relationship between self-esteem, affect, and leniency. By encouraging raters to focus on the performance related behaviors of ratees and not on how well the rater likes a ratee, leniency may be reduced and discrimination enhanced. In addition, by carefully defining performance dimensions, leniency may be minimized. Past research has found that the

relationship between self-esteem and leniency to be especially salient when performance dimension are ambiguous (Farh & Dobbins, 1989).

Organizations may be able to more directly control the effects of self-efficacy on leniency. In the current study, greater self-efficacy was related to increased leniency and less discrimination among dimensions. Self-efficacy was closely related to self-esteem, and it may be that in the current study that lenient ratings occurred when high self-esteem individuals were also high in efficacy for the evaluation task. Previous research has stressed the importance of having raters who understand their part in the evaluation process is to reducing leniency in ratings (Napier & Latham, 1986). While this is true, there may be some indication that when some individuals are both high in self-esteem and high in self-efficacy, they may be inclined to provide more lenient ratings.

While organizations may be limited to the degree of direct control they have over the influence of individual variables, processes such as rater training may help mitigate an potential negative effects of such variables on the evaluation process.

Strengths and Limitations of Current Study

The current study attempted to meet two primary goals: (1) to contribute a unique perspective to the literature by examining differences between self-rater and other-rater in terms of ratings given, and (2) to examine individual and contextual variables that influence rating behaviors associated with self and other-rater roles. While not all hypotheses were supported, general support was found for a model of rating behavior that

includes the effects of rating purpose, rating role, and individual and contextual variables on rating behavior.

Study design. A primary strength of the current study was in its design. Past research has limited its study of rating behavior and rater roles to comparison of ratings of self with those received from other sources. The design of the current study was unique in allowing for the examination of differences between self and other rating behaviors by comparing those given to self with those given to others. The inclusion of this unique perspective contributed to the literature by confirming that the same pattern of differences between self and other-rater behavior might be found regardless of whether comparisons are made at the level of self with those received or self with those given.

An additional design strength involves the collection of self-ratings from both those serving in supervisory and non-supervisory roles. Previous research has not examined self-ratings from both perspectives within the context of a single study. Such comparison reinforced that self-ratings under numerous conditions may be expected to differ from those associated with other-rater roles.

While limited previous research (e.g., Cleveland et. al, 2000) has examined the relationship between rating behavior and contextual variables, this relationship has not been examined in the context of different rater roles. In addition, research has not systematically examined the influence of both individual and contextual variables on rating behavior. The design of the current study permitted the systematic examination of

how individual and contextual variables might differently influence rating behavior depending on the role of the rater. While results were mixed, they support the importance that individual characteristics and contextual characteristics may have to rating behavior associated with different rater roles.

Finally, the design allowed for rating purpose (administrative or developmental) to be strictly controlled for. Had an organizational setting been used for the current study, the same degree of control may not have been possible. Often, organizations do not specifically indicate for what purposes ratings will be used (London & Smithers, 1995).

Despite the strengths of the current study, several possible weaknesses also deserve discussion. One possible weakness in the study's design may be that it was not fully crossed. Participants who served as group members in the study served in the roles of self-rater, peer-rater, and subordinate-rater (subordinate rating a supervisor). Participants who served as supervisor members served in the roles of self-rater and supervisor-rater (supervisor rating a subordinate). Because not all participants engaged in all rating roles, some statistical comparisons could not be made.

An additional weakness of the current study involves artificiality of the evaluation task and the participant sample. As it was not possible to include organizational members as participants in the study university students were used. The experimental groups were together for only a brief time (approximately one hour). It is probable that

the experience of these groups cannot be likened to those that occur with groups in organizations in which individuals interact on a daily basis. It may be argued that some of the covariate variables of interest (for example, politics in rating) would have had more influence if examined in the context of an organization's actual performance evaluation process.

Lastly, a possible weakness of the current study may be that each covariate variable was measured only once. Results for some covariates may have differed had they been measured for each rater role. For example, affect was measured by assessing in general how well the rater liked ratees. If affect had been measured for each rater role, more specific information could be gained regarding how well the rater liked a peer they rated versus a supervisor versus themselves.

Future Research

Future research examining rater roles and rating behavior may benefit from expanding on some of the current study's findings. Rater role, rating purpose, and covariates were all found to be significantly related to rating behavior of self and other-raters. Each may be discussed in terms of future research.

The current study examined the roles of self-rater, peer-rater, subordinate-rater, and supervisor-rater for their relationship to rating behavior. As the popularity of multi-rater evaluation systems grows so do the number of rater roles included in the process. It is becoming more common for "customers" (those who receive some service from the

ratee) to evaluate the performance of those who provide a service. Research has not examined the relationship between rater roles that are external to the organization and the rating behavior associated with such roles. Future research will benefit from examining the rating behavior associated with rater roles that are based both within and outside the organization.

Administrative and developmental purposes were examined in the current study for their relationship to rating behavior. Often, raters are not given a specific purpose for rating. In addition, some evidence suggests that even when raters are told that ratings will be used for developmental purposes they believe they will have some influence on administrative decisions. Future research will benefit from examining rating behavior when no purpose is given, or when raters are given a purpose but doubt the actual purpose for rating.

The current study found both individual and contextual variables to be important to self and other rating behavior. Covariates were found to vary in their significance to rating behavior depending on rater role (self, peer, subordinate, supervisor). Based on this finding, future research would be encouraged to examine covariates at the level of rater role. While only a limited number of covariates were examined in the current study, future research would benefit from further examination of variables that may be related to rating behavior. Of particular importance for study may be those more "proximal" (i.e. closer) to the rater. Opportunity to observe performance and the quality of exchange

between rater and ratee are two "proximal" variables that might be considered in future research.

In addition, future research should continue to examine the relationship between self-esteem and self-efficacy, and their relationship to rating behavior. Self-efficacy was included in the current study as a contextual variable that measured the degree to which individuals understood and felt capable of participating in the evaluation process. Self-esteem was included in the current study as an individual characteristic that measured the degree to which individuals feel "worthy". Based on the results of regression analysis, self-esteem appeared to be a more significant contributor to rating behavior of self-raters, while self-efficacy was a more significant contributor to rating behavior of other-raters.

Recent research (i.e., Gardner & Pierce, 1998), has examined self-esteem and self-efficacy as being a part of the individual's "self-concept". Under this approach, self-efficacy (even task-specific self-efficacy) might be considered as an individual as opposed to contextual variable in future research.

While specific hypotheses were not made regarding experimental source (participated in task as group member or supervisor member) in the current study, some significant results were found for self-raters. Interestingly, when ratings were associated with an administrative purpose, supervisor self-raters were more lenient in their ratings than were group self-raters. In contrast, ratings associated with a developmental purpose for group self-raters were more lenient than were supervisor self-ratings. Future studies

would benefit from examining the rater's position in the organization, along with rater role.

The current study was successful in confirming that ratings given to self and those given to others share the same pattern of discrepancy found when ratings given to self are compared with those received from others. In addition, the current study was successful in identifying individual and organizational variables that have importance for understanding the relationships between rater roles and rating behaviors. Findings from the current study have relevance for organizational implementation of multi-rater evaluation processes and for individuals interested in pursuing research in the area of rating behavior. Future research will benefit from expanding on the findings reported here and from examining additional questions and concerns involving rating behavior.

REFERENCES

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

Ashford, S. J. (1989). Self-assessment in organizations: A literature review and integrative model. In Cummings L. L. , Staw B. M. (Eds.), Research in Organizational Behavior, 11, (pp. 133-174). Greenwich, CT: JAI Press.

Atwater, L. E., & Waldman, D. (1998). Accountability in 360 degree feedback. HRMagazine, 43, 96-104.

Atwater, L., & Yammarino, F. J. (1992). Personal attributes as predictors of superiors' and subordinates' perceptions of military academy leadership. Human Relations, 46, 645-668.

Atwater, L. E., & Yammarino, F. J. (1997). Self-other ratings agreement: A review and a model. Research in Personnel and Human Resources Management, 15, 121-174.

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

Bettenhausen, K. L., & Fedor, D. B. (1991). Constructing evaluation realities: How the dual roles of giving and receiving peer feedback affect acceptance of a peer evaluation system. In M. Peiperl (Chair), Peer evaluation revisited. Symposium conducted at the 51st annual meeting of the Academy of Management, Miami, FL.

Bettenhausen, K. L., & Fedor, D. B. (1997). Peer and upward appraisals: A comparison of their benefits and problems. Group and Organization Management, 22, 236-263.

Borman, W. C. (1974). The rating of individuals in organizations: An alternative approach. Organizational Behavior and Human Performance, 12, 105-124.

Campbell, D. J., & Lee, C. (1988). Self-appraisal in performance evaluation: Development versus evaluation. Academy of Management Review, 13, 302-314.

Cleveland, J. N., Murphy, K. R., & Tziner, A. (2000). Relationships between attitudes toward organizations and performance appraisal systems and rating behavior: A multi-national study. Manuscript submitted for publication.

Cleveland, J. N., Murphy, K. R., & Williams, R. E. (1989). Multiple uses of performance appraisal: Prevalence and correlates. Journal of Applied Psychology, 74, 130-135.

Coates, D. (1998). Don't tie 360 feedback to pay. Training, 35, 68-78.

Fahr, J., Cannella, A. A., & Bedeian, A. G. (1991). Peer ratings: The impact of purpose on rating quality and user acceptance. Group and Organizational Studies, 16, 367-386.

Fahr, J., & Dobbins, G. (1989). Effects of comparative performance information on the accuracy of self-ratings and agreement between self and supervisor ratings. Journal of Applied Psychology, 74, 606-610.

Gardner, D. G., Pierce, J. L. (1998). Self-esteem and self-efficacy within the organizational context: An empirical examination. Group and Organization Management, 23, 48-70.

Gioia, D. A., Longenecker, C. O. (1994). Delving into the dark side: The politics of executive appraisal. Organizational Dynamics, 22, 47-58.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. Personnel Psychology, 41, 43-59.

Holzbach, R. L. (1978). Rater bias in performance ratings: Superior-, self-, and peer ratings. Journal of Applied Psychology, 59, 638-642.

Jolly, J. P., Reynolds, T. J., & Slocum, J. W. (1988). Application of the means-end theoretic for understating the cognitive bases of performance appraisal. Organizational Behavior and Human Performance, 41, 153-179.

Kingstrom, P. O., & Mainstone, L. E. (1985). An investigation of rater-ratee acquaintance and rater bias. Academy of Management, 28, 641-653.

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

Landy, F. J., & Farr, J. L. (1983). The measurement of work performance: Methods, theory, and applications. New York: Academic Press.

Lawler, E. (1967). The multi-trait multi-rater approach to measuring managerial job performance. Journal of Applied Psychology, 51, 369-381.

Lee, M. & Byoungho, S. (1999). The agreement between self-and supervisor ratings: An investigation of leader-member exchange effects. International Journal of Management, 16, 77-88.

Levy, P. E. (1991). Self-appraisal and attributions: A test of a model. Journal of Management, 19, 51-62.

London, M., & Smither, J. W. (1995). Can multiple-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. Personnel Psychology, 48, 803-840.

London, M. & Smither, J. W. (1998). Can multiple-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications an directions for research. Personnel Psychology, 48, 803-840.

Locke, E. A., Frederick, E., Lee, C., & Bobko, P. (1984). Effect of self-efficacy, goals, and task strategies on task performance. Journal of Applied Psychology, 69, 241-251.

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

Litwin, G. H., & Stringer, R. A. (1968). Motivation and organizational climate. Boston: Harvard University Press.

Mabe, P., & West, S. (1982). Validity of self-evaluation of ability: A review and meta-analysis. Journal of Applied Psychology, 67, 280-296.

Mathieu, J. E., & Zajac, D. M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. Psychological Bulletin, 108, 171-194.

McEvoy, G. M., & Buller, P. F. (1987). User acceptance of peer appraisals in an industrial setting. Personnel Psychology, 40, 785-797.

Mount, M. K. (1984). Satisfaction with a performance appraisal system and appraisal discussion. Journal of Occupational Behavior, 5, 271-279.

Mowday, R. T., Porter, L. W., & Steers, R. M. (1982). Employee-organizational linkages. New York: Academic Press.

Murphy, K. R., & Cleveland, J. N. (1991). Performance appraisal: An organizational perspective. Boston: Allyn and Bacon.

Napier, N., & Latham, G. (1986). Outcome expectancies of people who conduct performance appraisals. Personnel Psychology, 39, 827-837.

Robinson, R. K., Fink, R. L., & Allen, B. M. (1996). The influence of organizational constituent groups on rater attitudes toward performance appraisal compliance. Public Personnel management, 25, 141-151.

Rotter, J. B.(1966). Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 80, 1-28.

Schrader, B. W., & Steiner, D. D. (1996). Common comparison standards: An approach to improving agreement between self and supervisory performance ratings. Journal of Applied Psychology, 81, 813-820.

Steele, R., & Ovalle, N. (1984). Self-appraisal based upon supervisory feedback. Personnel Psychology, 37, 667-685.

Steers, R. M., & Lee, T. W. (1983). Facilitating effective performance appraisals: The role of the employee commitment and organizational climate. In F. Landy, S. Zedeck, & J. N. Cleveland (Eds.). Performance measurement and theory. (pp. 75-88). Hillsdale, NJ: Lawrence Erlbaum Associates.

Thornton, G. C. (1968). The relationship between supervisory-and self-appraisals of executive performance. Personnel Psychology, 21, 441-455.

Thornton, G. C. (1980). Psychometric properties of self-appraisals of job performance. Personnel Psychology, 33, 263-271.

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

Tsui, A. S., & Barry, B. (1986). Interpersonal affect and rating errors. Academy of Management Journal, 29, 586-599.

Varma, A. Denisi, A.S, & Peters, L. H. (1996). Interpersonal affect and performance appraisal: A field study. Personnel Psychology, 49, 341-361.

Vecchio, R. P., & Gobdel, B. C. (1984). The vertical dyadic linkage mode of leadership: Problems and prospects. Organizational Behavior and Human Performance, 34, 5-20.

Waldman, D. A., Atwater, L. E., Antonioni, D. (1998). Has 360 feedback gone amok? Academy of Management Executive, 12, 86-94.

Weisband, S. & Atwater, L. (1999). Evaluating self and others in electronic and face-to-face groups. Journal of Applied Psychology, 84, 632-639.

Westerman, J.W., & Rosse, J. G. (1997). Reducing the threat of rater nonparticipation in 360-degree feedback systems: an exploratory examination of antecedents to participation in upward ratings. Group and Organization Management, 22, 288-300.

Wohlers, A. J., & London, M. (1989). Ratings of managerial characteristics: Evaluation difficulty, co-worker agreement, and self-awareness. Personnel Psychology, 42, 235-261.

Zedeck, S., & Cascio, W. F. (1982). Performance appraisal decisions as a function of rater training and purpose of the appraisal. Journal of Applied Psychology, 67, 752-758.

Table 1

Means and Standard Deviations for Demographic, Covariate, and Dependent Variables for Full, Group, and Supervisor Samples.

Variable	n	M	SD
	Full Sample		
Demographic Variables[a]			
Age	179	1.25	.46
Sex	179	1.51	.50
Year in School	179	1.37	.70
Experience Rating Others	179	4.68	1.42
Experience Rating Self	179	4.90	1.48
Covariate Variables[b]			
Self-esteem	179	5.91	.80
Locus-of-control	179	5.39	.59
Politics	179	4.43	.85
Affect	179	4.45	1.03
Self-efficacy	179	5.87	.77
Organizational Commitment	179	4.86	1.00
Dependent Variables[c]			
Self-rater Leniency	179	5.95	.87

(table continues)

Variable	n	M	SD
Self-rater Dimension Discrimination	179	.42	.33
Other-rater Leniency	179	5.51	.85
Other-rater Dimension Discrimination	179	.53	.23
	Group Sample		
Demographic Variables			
Age	129	1.29	.49
Sex	129	1.49	.50
Year in School	129	1.37	.66
Experience Rating Others	129	4.66	1.38
Experience Rating Self	129	4.85	1.50
Covariate Variables			
Self-esteem	129	5.98	.80
Locus-of-control	129	5.42	.57
Politics	129	4.48	.83
Affect	129	4.45	1.02

(table continues)

Variable	n	M	SD
Self-efficacy	129	5.89	.74
Organizational Commitment	129	4.86	.97
Dependent Variables[c]			
Self-rater Leniency	129	5.93	.89
Self-rater Dimension Discrimination	129	.42	.34
Peer-rater Leniency	129	5.60	.77
Peer-rater Dimension Discrimination	129	.52	.29
Subordinate-rater Leniency	128[d]	5.53	1.11
Subordinate-rater Dimension Discrimination	128[d]	.58	.31
Discrimination between Peer Ratees (ratings given to peers)	71[e]	.55	.53
Supervisor Sample			
Demographic Variables			
Age	50	1.16	.37
Sex	50	1.58	.50
Year in School	50	1.38	.81

(table continues)

Variable	n	M	SD
Experience Rating Others	50	4.70	1.54
Experience Rating Self	50	5.06	1.41
Covariate Variables			
Self-esteem	50	5.75	.80
Locus-of-control	50	5.30	.63
Politics	50	4.32	.91
Affect	50	4.34	1.06
Self-efficacy	50	5.83	.86
Organizational Commitment	50	4.87	1.07
Dependent Variables			
Self-rater Leniency	50	5.98	.84
Self-rater Dimension Discrimination	50	.41	.29
Supervisor-rater Leniency	50	5.40	.90
Supervisor-rater Dimension Discrimination	50	.49	.24
Discrimination between Subordinate Ratees (ratings given to subordinates)	50	.61	.61

(table continues)

[a] The "Age" measure was scored as follows: "1" = under 20 years of age; "2"= 20-25; "3"= 25-30; "4"=over 30 years of age. The "Sex" measure was scored as follows: "1"= male; "2" = female. The "Year in School" measure was scored as follows: "1"= freshman; "2" = sophomore, "3"= junior, "4"= senior, and "5"= other. Experience evaluating self and experience evaluating other were represented by a 1-7 point scale. Both measures were scored such that "7" = great amount of evaluation experience and "1"= little experience with evaluation.

[b] Covariate variables were measured using 1-7 point scales. All measures were scored such that a "7" represented a positive rating while a "1" represented a more negative rating.

[c] Leniency is represented by the mean level of performance rating associated with each rater role. The "performance rating" measure was scored so that "7"= outstanding performance and "1"= poor performance. Dimension discrimination is represented by the mean level of discrimination between performance dimensions associated with each rater role. Numbers closer to .00 indicate low levels of rating discrimination between performance dimensions, while those closer to 1.00 indicate greater levels of dimension discrimination discrimination. Self-rater refers to an individual evaluating his or her own performance. Other-rater refers to all raters other than self. Peer-rater refers to an individual evaluating another group member. Subordinate-rater refers to a subordinate evaluating a supervisor. Supervisor-rater refers to a supervisor evaluating a subordinate.

[d] One participant did not complete a rating of their supervisor.

[e] Some peer raters were not included because they rated only one other peer due to group size.

Table 2

Intercorrelations Between Demographic, Covariate, and Dependent Variables of Interest

Variable	1	2	3	4	5	6	7	8	9
1. Age									
2. Sex	-.17*								
3. Year	.61**	-.06							
4. Experience Rating Others	.06	.05	.04						
5. Experience Rating Self	.00	.01	-.03	.44**					
6. Self-esteem	.06	-.09	-.01	.10	.27**				
7. Locus-of-control	.01	.05	.07	.02	.14	.46**			
8. Rating Politics	-.04	-.16*	.000	.02	.09	.10	-.07		
9. Affect	.01	-.06	-.02	-.02	.15*	.14	.10	.02	
10. Self-efficacy	.00	.16*	.03	.19*	.17*	.34**	.28**	-.03	.08
11. Organizational Commitment	.04	.16*	.04	.07	-.00	.04	.09	.04	.05

(table continues)

Variable	1	2	3	4	5	6	7	8	9
				Full Sample (N=179)					
12. Self -rater Leniency	.-.01	.09	-.04	.13	.15*	.20**	-.04	.04	.11
13. Self-rater Dimension Discrimination	.05	-.03	.04	-.14	-.13	-.22**	.03	-.06	-.13
14. Other-rater Leniency	.05	-.09	-.09	.04	.10	.20**	.04	-.09	.36**
15. Other-rater Dimension Discrimination	-.01	.04	.01	-.02	-.20**	-.15	.01	.08	-.13
				Group Sample (n=129)					
16. Self-rater Leniency	.03	.03	-.04	.09	.16	.24**	.07	.08	.08
17. Self-rater Dimension Discrimination	.02	.03	.03	-.13	-.15	-.26**	-.06	-.07	-.07
18. Peer-rater Leniency	.02	-.06	-.03	.06	.09	.29**	.14	.05	.24*
19. Peer-rater Dimension Discrimination	.07	-.06	.06	-.04	-.14	-.26**	-.07	-.03	-.09
20. Subordinate-rater Leniency	.06	-.08	-.03	-.01	.15	.20*	.07	-.18*	.24*
21. Subordinate-rater Dimension Discrimination	-.04	.10	-.01	-.06	-.15	-.12	.01	.12	-.07
				Supervisor Sample (n=50)					
22. Self-rater Leniency	-.15	.26	-.04	.23	.12	.01	-.30*	-.06	.19
23. Self-rater Dimension Discrimination	.14	-.22	.08	-.16	-.05	-.12	.30*	-.03	-.30*

(table continues)

Variable	1	2	3	4	5	6	7	8	9
24. Supervisor-rater Leniency	.01	-.08	.03	.07	.02	.01	-.16	-.11	.53**
25. Supervisor-rater Dimension Discrimination	-.14	.10	-.05	.07	-.23	.02	.09	.11	-.20
Discrimination Between Ratees									
26. Discrimination Between Peer Ratees	.01	.25	.08	-.18	-.09	-.21	.07	-.16	-.23
27. Discrimination Between Subordinate Ratees	.10	.06	.21	-.23	-.18	-.22	.14	-.17	-.12

(table continues)

Variable	10	11	12	13	14	15	16	17
11. Organizational Commitment	.06							
Full Sample (N=179)								
12. Self-rater Leniency	.20**	.17*						
13. Self-rater Dimension Discrimination	-.17*	.06	-.47					
14. Other-rater Leniency	.23**	.06	.47**	-.37**				
15. Other-rater Dimension Discrimination	-.06	-.07	-.16*	.28**	-.37**			
Group Sample (n=129)								
16. Self-rater Leniency	.18*	.18*						
17. Self-rater Dimension Discrimination	-.18*	.07	-.43**					
18. Peer-rater Leniency	.25**	.09	.42**	-.28**	.83**	-.31**	.42**	-.28**
19. Peer-rater Dimension Discrimination	-.15	-.06	-.16	.37**	-.31**	.75**	-.16	.37**
20. Subordinate-rater Leniency	.22*	-.01	.46**	-.30**	.92**	-.45**	.46**	-.30**
21. Subordinate-rater Dimension Discrimination	.01	.02	-.20*	.12	-.38**	.79**	-.20**	.12
Supervisor Sample (n=50)								
22. Self-rater Leniency	.24	.14						
23. Self-rater Dimension Discrimination	-.16	.05	-.60**					

(table continues)

Variable	10	11	12	13	14	15	16	17
24. Supervisor-rater Leniency	.15	.11	.42**	-.50**				
25. Supervisor-rater Dimension Discrimination	-.01	-.16	.04	.18	-.24			
Discrimination Between Rates								
26. Discrimination Between Peer Rates	-.14	.11	-.11	.14	-.30*	-.02	-.11	.14
27. Discrimination Between Subordinate Rates	-.24	-.13	-.18	.31*	-.38**	.13		

(table continues)

Variable	18	19	20	21	22	23	24
				Group Sample (n=129)			
22. Peer-rater Dimension Discrimination	-.35**						
23. Subordinate-rater Leniency	.53**	-.21*					
24. Subordinate-rater Dimension Discrimination	-.12	.18*	-.48**				
				Supervisor Sample (n=50)			
22. Self-rater Leniency							
23. Self-rater Dimension Discrimination							
24. Supervisor-rater Leniency	.15	.11	.42**	-.50**			
25. Supervisor-rater Dimension Discrimination	-.01	-.16	.04	.18	-.24		
				Discrimination Between Ratees			
26. Discrimination Between Peer Ratees	.-.44**	-.09	-.15	.03			
27. Discrimination Between Subordinate Ratees					-.18	.31*	-.38**

(table continues)

Variable	25	26	27
	Discrimination Between Ratees		
26. Discrimination Between Peer Ratees			
27. Discrimination Between Subordinate Ratees	.13		

Note. Some correlations between dependent variables were not applicable and are represented by empty cells.

*p<.05. **p<.01.

Table 3

Univariate F-test Results: Main effects and interactions (including covariates) for the Full Sample.

Measure	Effect	F (df)	Significance of F	Eta Squared
Leniency				
	Rater Role	61.22 (1,175)	.00	.26
	Rating Purpose	14.92 (1,175)	.00	.08
	Rater Role X Rating Purpose	50.94 (1,175)	.00	.23
	Rater Role X Self-esteem	.09 (1,177)	.77	.00
	Rater Role X Locus-of-Control	2.61 (1,177)	.11	.00
Discrimination Between Performance Dimensions				
	Rater Role	14.43 (1,175)	.00	.08
	Rating Purpose	.36 (1,175)	.56	.00
	Rating Role X Rating Purpose	9.09 (1,175)	.00	.05

Table 4

Means and Standard Deviations for Rating Purpose by Rater Role and Rating Behavior in the Full Sample

Measure	n	M	SD
Leniency			
Administrative Purpose			
Self Rater	85	6.37	.79
Other Rater	85	5.48	.80
Total		5.96	.09[a]
Developmental Purpose			
Self Rater	94	5.56	.76
Other Rater	94	5.53	.90
Total		5.50	.08[a]
Dimension Discrimination			
Administrative Purpose			
Self Rater	85	.37	.35
Other Rater	85	.56	.25
Total		.49	.03[a]

(table continues)

Measure	n	M	SD
Developmental Purpose			
Self Rater	94	.46	.30
Other Rater	94	.51	.22
Total		.48	.03[a]

[a]Refers to Standard Error.

Table 5

Univariate F-test Results: Main effects and Interactions (including covariates) for the Group Sample.

Measure	Effect	F (df)	Significance of F	Eta Squared
Leniency				
	Rater Role	17.03 (2,125)	.00	.21
	Rating Purpose	1.40 (1,126)	.24	.01
	Rater Role X Rating Purpose	21.24 (2,125)	.00	.25
	Rater Role X Self-esteem	.30 (2,125)	.74	.01
	Rater Role X Locus-of-Control	.74 (2,125)	.47	.00
Discrimination Between Performance Dimensions				
	Rater Role	9.15 (2,125)	.00	.13
	Rating Purpose	.08 (1,126)	.77	.00

(table continues)

Measure	Effect	F (df)	Significance of F	Eta Squared
	Rating Role X Rating Purpose	5.60 (2,125)	.00	.08
Discrimination Between Ratees (ratings given to peers)				
	Rating Purpose	1.0 (1,69)	.31	.02

Table 6

Means and Standard Deviations for Rating Purpose by Rater Role and Rating Behavior in the Group Sample

Rater Role	n	M	SD
Leniency			
Administrative Purpose			
Self Rater	61	6.28	.86
Peer Rater	61	5.46	.82
Subordinate Rater	60[a]	5.54	1.00
Total		5.77	.10[b]
Developmental Purpose			
Self Rater	68	5.62	.79
Peer Rater	68	5.71	.73
Subordinate Rater	68	5.51	1.20
Total		5.61	.09[b]

(table continues)

Rater Role	n	M	SD
Dimension Discrimination			
Administrative Purpose			
Self Rater	61	.39	.36
Peer Rater	61	.57	.29
Subordinate Rater	60[a]	.55	.26
Total		.50	.03
Developmental Purpose			
Self Rater	68	.45	.32
Peer Rater	68	.46	.27
Subordinate Rater	68	.61	.34
Total		.51	.03[b]

Note. Subordinate refers to a subordinate rating a supervisor.
[a]One participant did not complete a rating of their supervisor.
[b]Refers to Standard Error .

Table 7

Univariate F-test Results: Main effects and Interactions (including covariates) for the Supervisor Sample.

Measure	Effect	F (df)	Significance of F	Eta Squared
Leniency				
	Rater Role	30.16 (1,48)	.00	.39
	Rating Purpose	12.07 (1,48)	.00	.20
	Rater Role X Rating Purpose	22.79 (1,48)	.00	.32
	Rater Role X Self-esteem	.74 (1,48)	.39	.02
	Rater Role X Locus-of-Control	1.15 (1,177)	.29	.02
Discrimination Between Performance Dimensions				
	Rater Role	3.48 (1,48)	.07	.07
	Rating Purpose	.17 (1,48)	.67	.00
	Rating Role X Rating Purpose	7.96 (1,48)	.01	.14

(table continues)

Measure	Effect	F (df)	Significance of F	Eta Squared
Discrimination Between Ratees (ratings given to subordinates)				
	Rating Purpose	1.51 (1,48)	.23	.03

Table 8

Means and Standard Deviations for Rating Purpose by Rater Role and Rating Behavior in the Supervisor Sample

Measure	n	M	SD
Leniency			
Administrative Purpose			
Self Rater	24	6.60	.50
Supervisor Rater	24	5.46	.80
Total		6.03	.14
Developmental Purpose			
Self Rater	26	5.42	.68
Supervisor Rater	26	5.33	.99
Total		5.38	.13
Dimension Discrimination			
Administrative Purpose			
Self Rater	24	.33	.31
Supervisor Rater	24	.55	.29
Total		.44	.04
Developmental Purpose			
Self Rater	26	.49	.24

(table continues)

Measure	n	M	SD
Supervisor Rater	26	.44	.19
Total		.47	.04

Note. Supervisor rater refers to a supervisor rating a subordinate.

Table 9

Summary of Multiple Regression Analyses for Covariates Predicting Rating Behavior for the Full Sample of Participants (N=179)

Measure	Covariate	B	SE B	Standardized Beta
Rating Leniency				
Self Rater				
Step 1	Self-esteem	.22	.08	.20**
Step 2	Self-esteem	.21	.08	.20**
	Organizational Commitment	.14	.06	.16*
Other Rater				
Step 1	Affect	.30	.06	.35**
Step 2	Affect	.28	.06	.34**
	Self-efficacy	.22	.08	.20**
Dimension Discrimination				
Self Rater				
Step 1	Self-esteem	-9.10	.03	-.22**

Note. $R^2 = .04$ for Step 1, self rater leniency. R^2 Change = .03, Step 2, self rater leniency ($p<.05$). $R^2 = .13$ for Step 1, other rater leniency. R^2 Change= .04, Step 2, other rater leniency. $R^2 = .05$ for Step 1, self-rater discrimination.

Table 10

ANCOVA Results: Comparing changes in F statistics when controlling for covariates in the Full Sample.

Measure	Effect	F (df)	Significance of F	Eta Squared
Leniency				
	Rater Role			
	Covariates in analysis:	61.22 (1,175)	.00	.26
	Covariates controlled for:	1.55 (1,175)	.21	.01
	Rating Purpose			
	Covariates in analysis:	14.92 (1,175)	.00	.08
	Covariates controlled for:	13.84 (1,172)	.00	.07
	Rater Role X Rating Purpose			
	Covariates in analysis:	50.94 (1,175)	.00	.23
	Covariates controlled for:	61.78 (1,172)	.00	.26

(table continues)

Measure	Effect	F (df)	Significance of F	Eta Squared
Discrimination Between Performance Dimensions				
	Rater Role			
	Covariates in analysis:	14.43 (1,175)	.00	.08
	Covariates controlled for:	.49 (1,172)	.49	.00
	Rating Purpose			
	Covariates in analysis:	.36 (1,175)	.56	.00
	Covariates controlled for:	.51 (1,172)	.47	.00
	Rating Role X Rating Purpose			
	Covariates in analysis:	9.09 (1,175)	.00	.05
	Covariates controlled for:	6.70 (1,172)	.01	.04

Table 11

Summary of Multiple Regression Analyses for Covariates Predicting Rating Behavior for the Group Sample of Participants (N=129)

Measure	Covariate	B	SE B	Standardized Beta
Rating Leniency				
Self Rater				
Step 1	Self-esteem	.28	.10	.25**
Step 2	Self-esteem	.28	.09	.25**
	Organizational Commitment	.16	.08	.18*
Peer Rater				
Step 1	Self-esteem	.29	.08	.29**
Step 2	Self-esteem	.25	.08	.26**
	Affect	.15	.07	.20*
Subordinate Rater				
Step 1	Affect	.27	.10	.24**
Step 2	Affect	.25	.09	.22**
	Self-efficacy	.31	.13	.20*
Step 3	Affect	.24	.09	.22**
	Self-efficacy	.33	.13	.22**
	Politics	-.27	.11	-.20*

(table continues)

Measure	Covariate	B	SE B	Standardized Beta
Dimension Discrimination				
Self Rater				
Step 1	Self-esteem	-.11	.04	-.26**
Peer Rater				
Step 1	Self-esteem	-.10	.03	-.26**

Note. For self rater leniency: R^2 = .06, Step 1; R^2 Change = .03 (p<.05), Step 2.
For peer rater leniency: R^2 = .09, Step 1; R^2 Change = .04 (p<.05), Step 2.
For subordinate rater leniency: R^2= .06, Step 1; R^2 Change = .04 (p<.01), Step 2; R^2 Change = .04 (p<.05), Step 3.
For self rater discrimination: R^2 = .07.
For peer rater discrimination: R^2=.07.

Table 12

ANCOVA Results: Comparing changes in F statistics when controlling for covariates in the Group Sample.

Measure	Effect	F (df)	Significance of F	Eta Squared
Leniency				
	Rater Role			
	Covariates in analysis:	17.03 (2,125)	.00	.21
	Covariates controlled for:	.37 (2,121)	.69	.00
	Rating Purpose			
	Covariates in analysis:	1.40 (1,126)	.24	.01
	Covariates controlled for:	2.11 (1,122)	.15	.02
	Rater Role X Rating Purpose			
	Covariates in analysis:	21.24 (2,125)	.00	.25
	Covariates controlled for:	22.26 (2,121)	.00	.27

(table continues)

Measure	Effect	F (df)	Significance of F	Eta Squared
Discrimination Between Performance Dimensions				
	Rater Role			
	Covariates in analysis:	9.15 (2,125)	.00	.13
	Covariates controlled for:	1.68 (2,121)	.19	.00
	Rating Purpose			
	Covariates in analysis:	.08 (1,126)	.77	.00
	Covariates controlled for:	.28 (1,122)	.60	.00
	Rating Role X Rating Purpose			
	Covariates in analysis:	5.60 (2,125)	.00	.08
	Covariates controlled for:	5.43 (2,121)	.01	.08

Table 13

Summary of Multiple Regression Analyses for Covariates Predicting Rating Behavior for the Supervisor Sample of Participants (N=50)

Measure	Covariate	B	SE B	Standardized Beta
Rating Leniency				
Self Rater				
Step 1	Locus-of-control	-.40	.19	-.30**
Supervisor Rater				
Step 1	Affect	.45	.10	.53**
Dimension Discrimination				
Self Rater				
Step 1	Locus-of-control	.14	.06	.30*
Step 2	Locus-of-control	.14	.06	.32**
	Affect	-8.49	.04	-.31*

Note. For self rater leniency: R^2 = .09 for Step 1. For supervisor rater leniency: R^2 = .29. For self rater discrimination: R^2 = .30, Step 1; R^2 Change =.01, Step 2 ($p<.05$).

Table 14

ANCOVA Results: Comparing changes in F statistics when controlling for covariates in the Supervisor Sample.

Measure	Effect	F (df)	Significance of F	Eta Squared
Leniency				
	Rater Role			
	Covariates in analysis:	30.16 (1,48)	.00	.39
	Covariates controlled for:	1.81 (1,46)	.19	.04
	Rating Purpose			
	Covariates in analysis:	12.07 (1,48)	.00	.20
	Covariates controlled for:	13.13 (1,46)	.00	.22
	Rater Role X Rating Purpose			
	Covariates in analysis:	22.79 (1,48)	.00	.32
	Covariates controlled for:	24.23 (1,46)	.00	.35

(table continues)

Measure	Effect	F (df)	Significance of F	Eta Squared
Discrimination Between Performance Dimensions				
	Rater Role			
	Covariates in analysis:	3.48 (1,48)	.07	.07
	Covariates controlled for:	.17 (1,46)	.67	.00
	Rating Purpose			
	Covariates in analysis:	.17 (1,48)	.67	.00
	Covariates controlled for:	.00 (1,46)	.97	.00
	Rating Role X Rating Purpose			
	Covariates in analysis:	7.96 (1,48)	.01	.14
	Covariates controlled for:	6.43 (1,46)	.02	.12

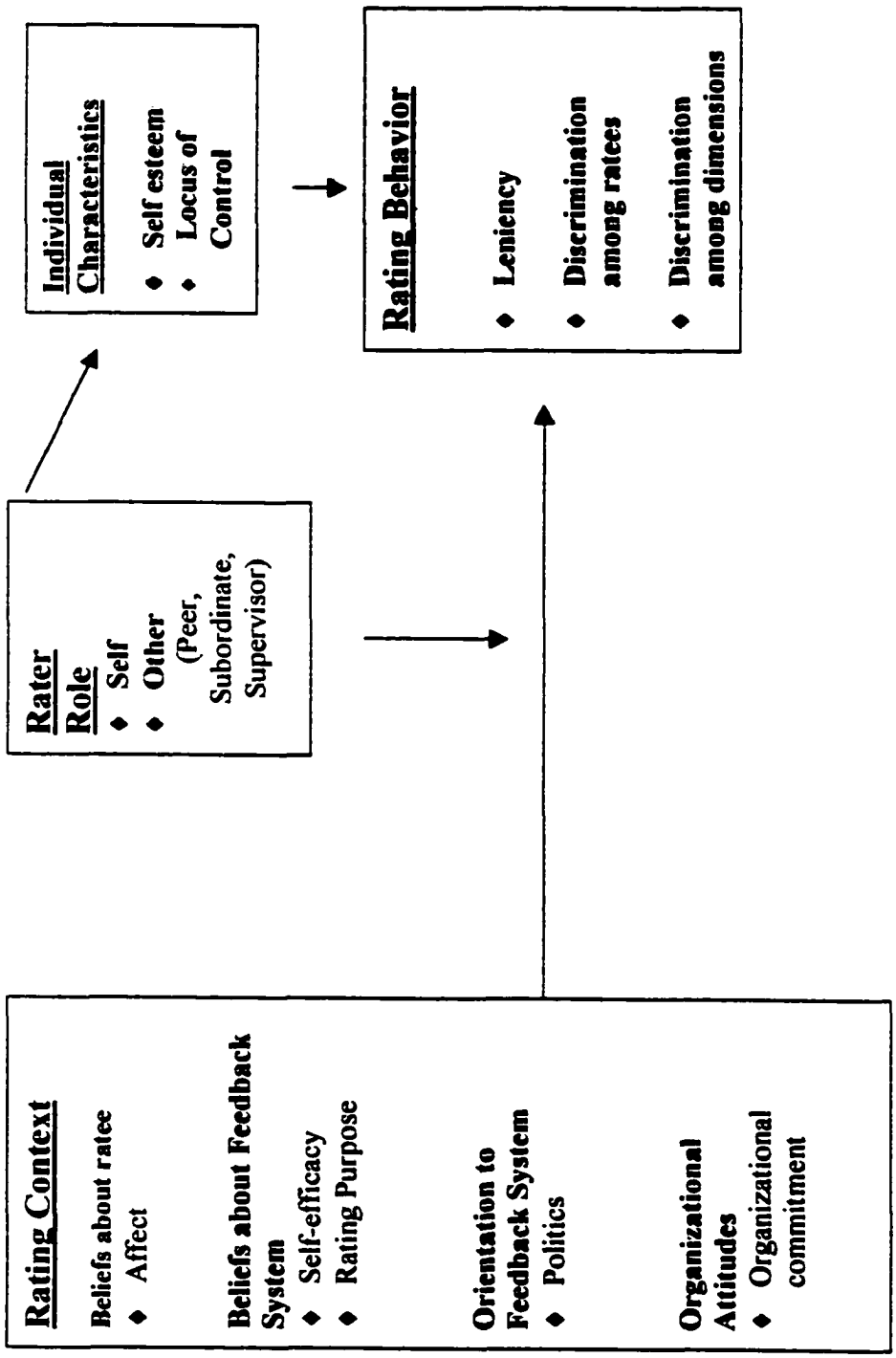

Figure 1. The influence of individual and contextual variables on the rating behavior of self, peer, subordinate and supervisor rater.

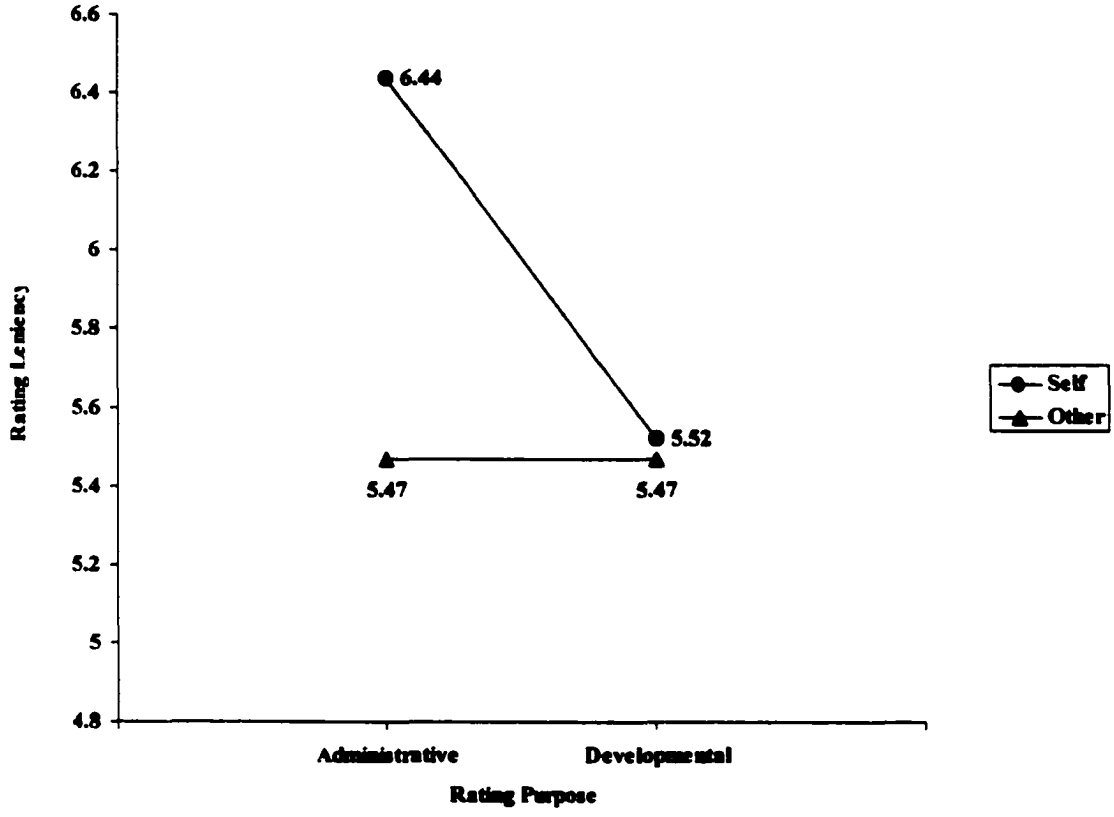

Figure 2. Interaction of rater role and rating purpose on rating leniency for the full sample.

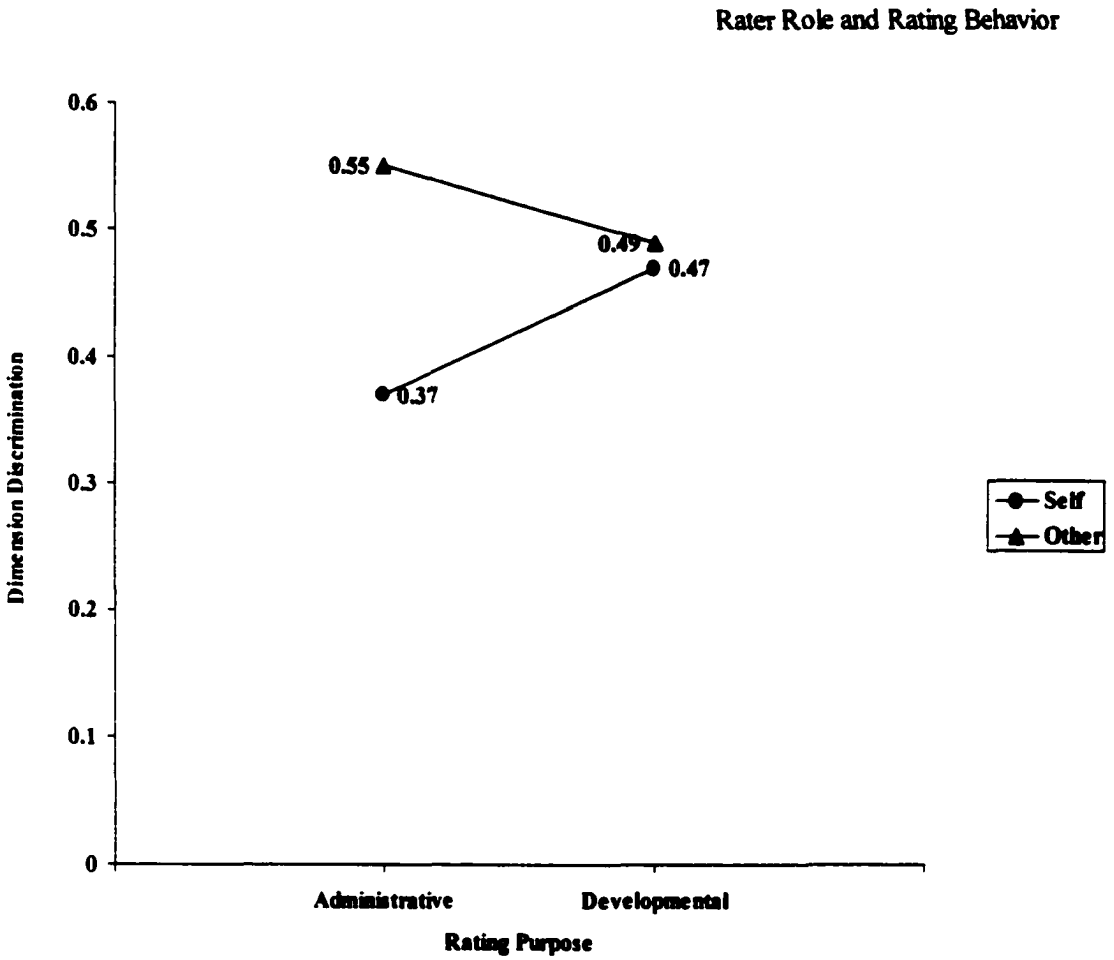

Figure 3. Interaction of rater role and rating purpose on discrimination between performance dimensions for the full sample.

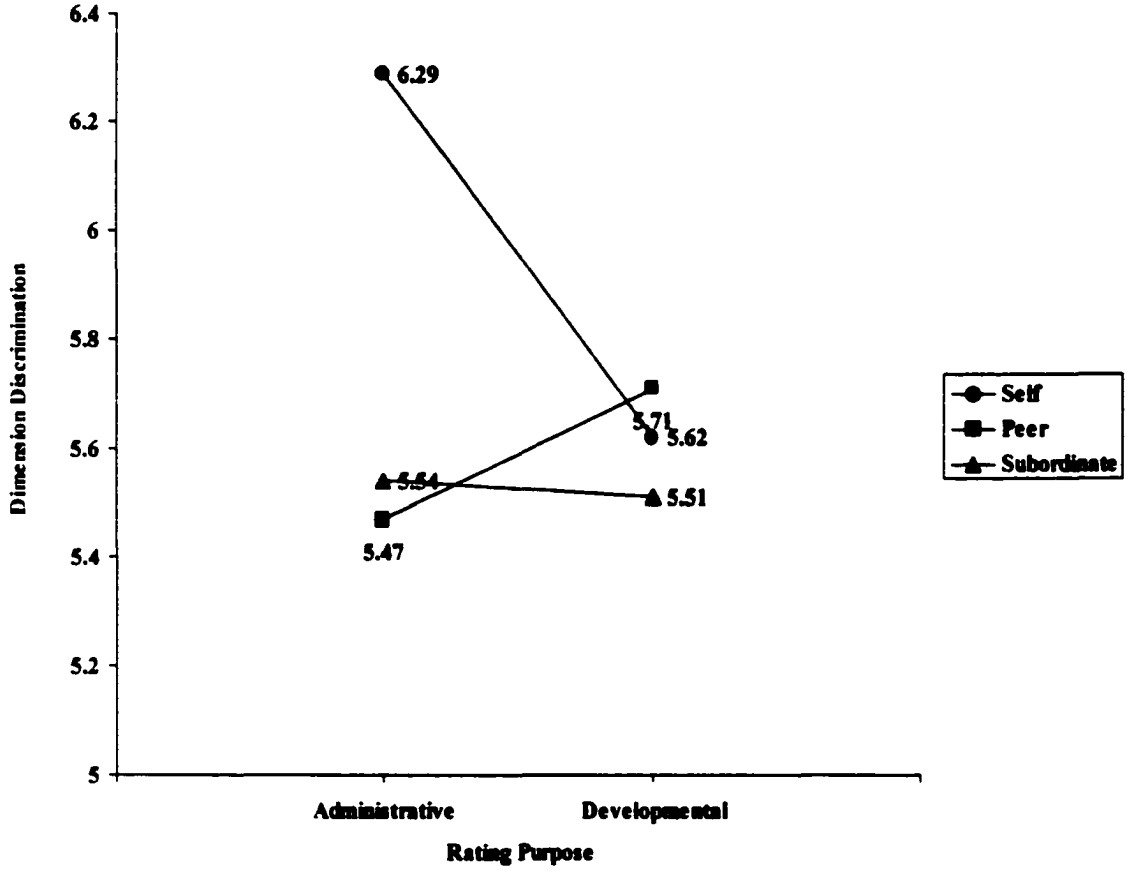

Figure 4. Interaction of rater role and rating purpose on rating leniency for the group sample.

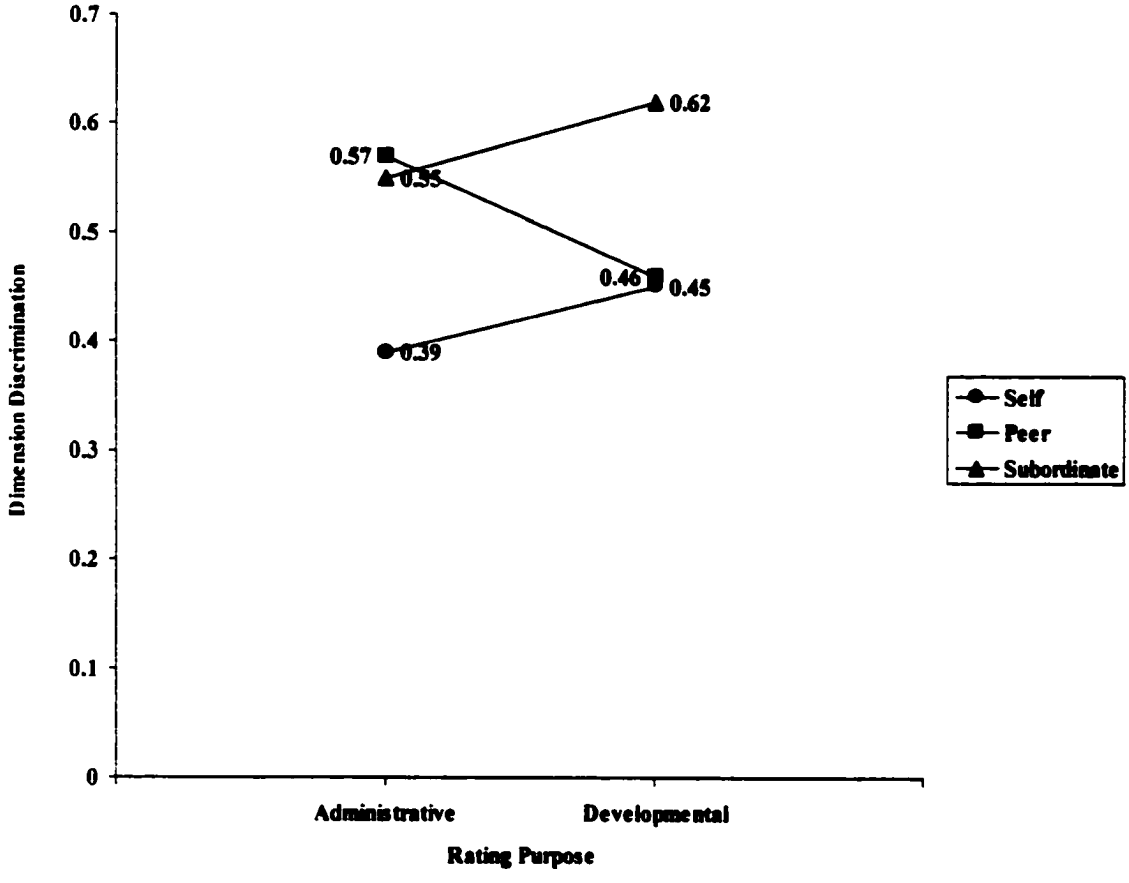

Figure 5. Interaction of rater role and rating purpose on dimension discrimination for the group sample.

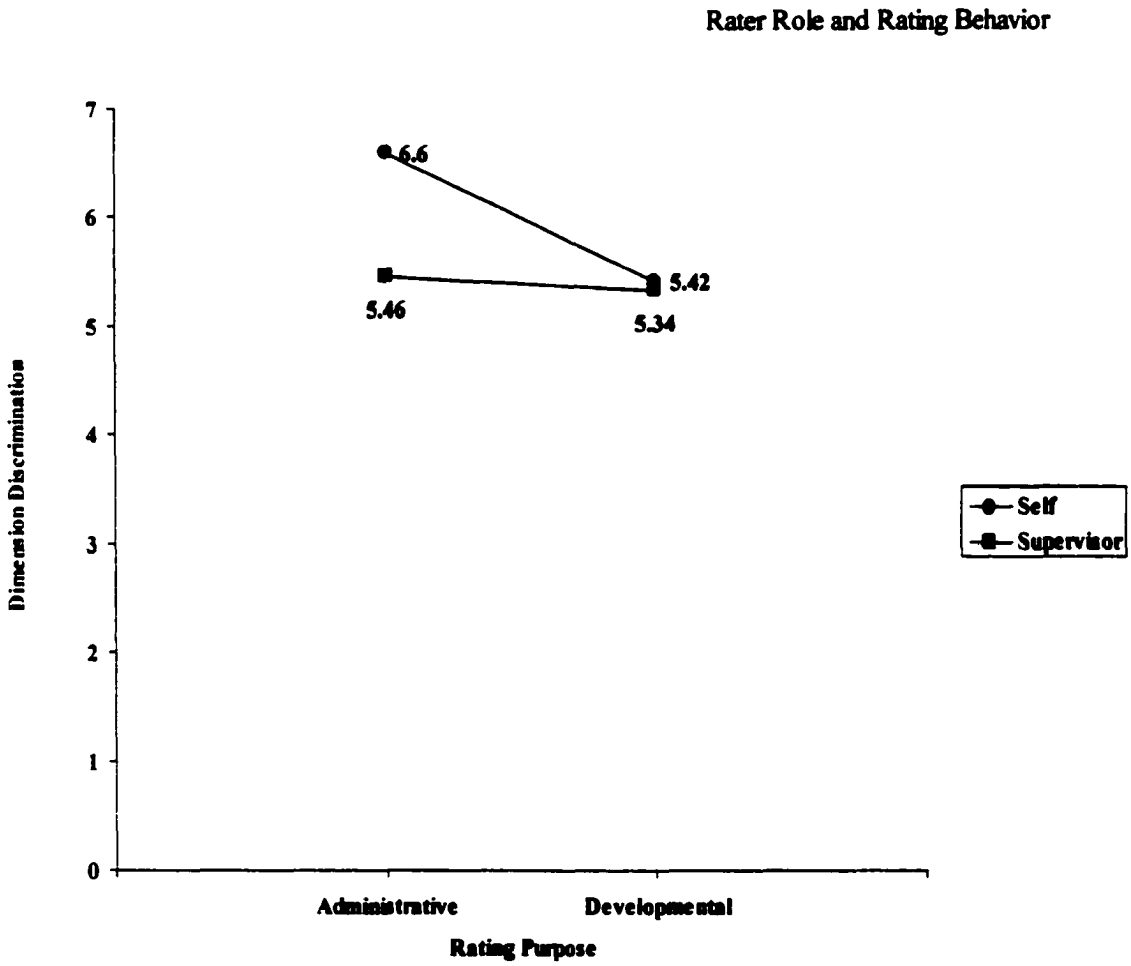

Figure 6. Interaction of rater role and rating purpose on rating leniency for the supervisor sample.

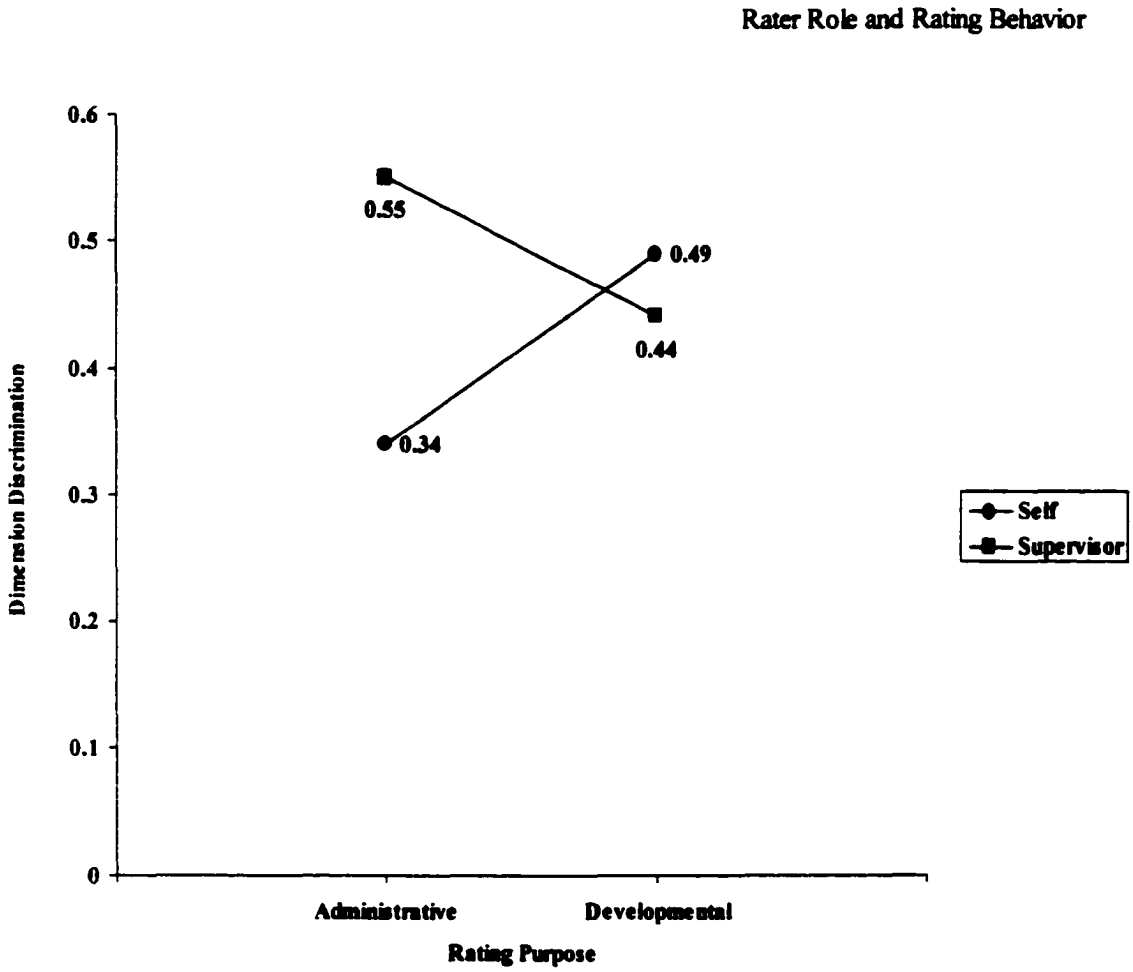

Figure 7. Interaction of rater role and rating purpose on dimension discrimination for the supervisor sample.

Appendix A

Task Instructions to Supervisor

You will serve as a supervisor for the group members involved in the idea generation task. As a supervisor, you will be responsible for assisting group members as they participate in the group task. You are required to sit outside of the group, but may interact with the group in a leadership capacity. Below are listed some instructions for your role as supervisor.

- Select a group member to record the group's ideas as they generate them and provide them with paper for recording purposes. The experimenter will provide paper.
- Inform group members that you will be tracking the time and will alert them when they have 2 minutes left to complete the task, and then will alert them when time is up. The experimenter will prompt you as to what time is left.
- As the group engages in the task, observe the members but do not contribute to their discussion.
- If the group appears to be "stuck", you may direct them with brief comments. For instance, you might say "Is there anything additional anyone wanted to add?" or "Is there anything else that needs to be considered?"

- When time is up, you will instruct group members to finish their final thoughts. At that time select a group member to present the group's conclusions to both you and the experimenter.
- You will listen to the group's presentation and then will proceed with the study as instructed by the experimenter.

Appendix B

Instructions and Rating Forms for Group Members

Instructions for Group Members

On the following pages you will find several rating forms and questionnaires. The rating forms will be used for you to evaluate the performance of each of your group members, your supervisor, and your own performance. The questionnaires ask questions about you, the people you rated, and about the rating process. Use your experience in the group task to guide you as you complete the questionnaires. It is important to complete the measures as honestly and accurately as you can. Please do not put names on any of the measures so that your responses can remain anonymous.

You may now turn to the ratings and questionnaires.

Rating Forms for Group Members

Instructions.

Use the following scales to rate the performance of a **group member.** Think of this individual's performance on the group task and rate their performance as accurately as possible using the scales below. Use the blank space below to indicate which group member you are rating by writing in the **letter** that appears on their nametag. Do not write down your name or the name of the group member.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your group member's performance.

Group member: _____

Task Proficiency (how well your group member performed the central elements of tasks that he or she was responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well your group member supported peers and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which your group member committed him or her self to job tasks, works at high intensity levels, and kept working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which your group member was able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which your group member followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your group member's performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Instructions.

Use the following scales to rate the performance of a **group member**. Think of this individual's performance on the group task and rate their performance as accurately as possible using the scales below. Use the blank space below to indicate which group member you are rating by writing in the **letter** that appears on their nametag. Do not write down your name or the name of the group member.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your group member's performance.

Group member: _____

Task Proficiency (how well your group member performed the central elements of tasks that he or she was responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well your group member supported peers and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which your group member committed him or her self to job tasks, works at high intensity levels, and keeps working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which your group member was able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which your group member followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your group member's performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Instructions.

Use the following scales to rate the performance of a group **member.** Think of this individual's performance on the group task and rate their performance as accurately as possible using the scales below. Do not write down your name or the name of the group member.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your group member's performance.

Group member: _____

Task Proficiency (how well your group member performed the central elements of tasks that he

or she is responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well your group member supported peers and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which your group member committed him or her self to job tasks, worked at high intensity levels, and kept working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which your group member was able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which your group member followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your group member's performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Instructions.

Use the following scales to rate the performance of your **supervisor**. Think of this individual's performance and rate their performance as accurately as possible using the scales below. Do not write down your name or the name of the supervisor.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your supervisor's performance.

Task Proficiency (how well your supervisor performed the central elements of tasks that he or she is responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well your supervisor supported the group and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which your supervisor committed him or her self to job tasks, worked at high intensity levels, and kept working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which your supervisor was able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which your supervisor followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your supervisor's performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Instructions.

Use the following scales to rate **your own performance** on the group task. Think of your performance and rate your performance as accurately as possible using the scales below. Do not write down your name.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your performance.

Task Proficiency (how well you performed the central elements of tasks that you were responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well you supported peers and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which you committed yourself to job tasks, worked at high intensity levels, and kept working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which you were able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which you followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Appendix C

Instructions and Rating Forms for Supervisors

Instructions for Supervisors

On the following pages you will find several rating forms and questionnaires. The rating forms will be used for you to evaluate the performance of each group member you supervised and your own performance. The questionnaires ask questions about you, the people you rated, and about the rating process. Use your experience in the group task to guide you as you complete the questionnaires. It is important to complete the measures as honestly and accurately as you can. Please do not put names on any of the measures so that your responses can remain anonymous.

You may now turn to the ratings and questionnaires.

Rating Forms for Supervisors

Instructions.

Use the following scales to rate the performance of a **group member.** Think of this individual's performance on the group task and rate their performance as accurately as possible using the scales below. Use the blank space below to indicate which group member you are rating by writing in the **letter** that appears on their nametag. Do not write down your name or the name of the group member.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your group member's performance.

Group member: _____

Task Proficiency (how well your group member performed the central elements of tasks that he or she was responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well your group member supported peers and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which your group member committed him or her self to job tasks, works at high intensity levels, and kept working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which your group member was able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which your group member followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your group member's performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Instructions.

Use the following scales to rate the performance of a **group member**. Think of this individual's performance on the group task and rate their performance as accurately as possible using the scales below. Use the blank space below to indicate which group member you are rating by writing in the **letter** that appears on their nametag. Do not write down your name or the name of the group member.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your group member's performance.

Group member: _____

Task Proficiency (how well your group member performed the central elements of tasks that he or she was responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well your group member supported peers and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which your group member committed him or her self to job tasks, works at high intensity levels, and keeps working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which your group member was able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which your group member followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your group member's performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Instructions.

Use the following scales to rate the performance of a group **member.** Think of this individual's performance on the group task and rate their performance as accurately as possible using the scales below. Do not write down your name or the name of the group member.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your group member's performance.

Group member: _____

Task Proficiency (how well your group member performed the central elements of tasks that he or she is responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well your group member supported peers and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which your group member committed him or her self to job tasks, worked at high intensity levels, and kept working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which your group member was able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which your group member followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your group member's performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Instructions.

Use the following scales to rate **your own performance** on the group task. Think of your performance and rate your performance as accurately as possible using the scales below. Do not write down your name.

Examples of each performance dimension are provided in parentheses. Circle the number that appropriately describes your performance.

Task Proficiency (how well you performed the central elements of tasks that you were responsible for)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Facilitating Peer and Team Performance (how well you supported peers and helped them with problems)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Demonstrating Effort (the degree to which you committed yourself to job tasks, worked at high intensity levels, and kept working under adverse conditions)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Oral Communication (the degree to which you were able to present material orally)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Following Policies and Procedures (the degree to which you followed rules, policies and procedures)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Overall Performance (how would you rate your performance overall)

Poor			Average		Outstanding	
1	2	3	4	5	6	7

Appendix D

Self-esteem

Instructions. This section contains several statements about how you feel about yourself. Indicate the extent to which you agree or disagree with each statement, using the scale below:

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree

	Rating
1. I feel that I'm a person of worth, at least on an equal basis with others.	_____
2. I feel that I have a number of good qualities.	_____
3. All in all, I am inclined to feel that I am a failure.	_____
4. I am able to do things as well as other people.	_____
5. I feel I do not have much to be proud of.	_____
6. I take a positive attitude toward myself.	_____
7. On the whole, I am satisfied with myself.	_____

8. I wish I could have more respect for myself. _____

9. I certainly feel useless at times. _____

10. At times I think I am no good at all. _____

Appendix E

Locus of Control

Instructions. This section contains several statements about how you feel about yourself. Indicate the extent to which you agree or disagree with each statement, using the scale below:

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree

	Rating
1. School is what I make of it.	____
2. At school I can accomplish whatever I set out to accomplish.	____
3. If I know what I want out of school, I can find a way to get it.	____
4. If I were unhappy with a decision made by my professor, I would do something about it .	____
5. Getting into the school I want is a matter of luck.	____
6. Getting good grades is primarily a matter of good fortune.	____
7. I am capable of doing well at school if I make the effort.	____

8. In order to get into a really good school I would need to have family members of friends in high places. _____

9. I believe that getting academic awards is usually a matter of good fortune. _____

10. When it comes to getting into a really good school, who I know is more important than what I can do. _____

11. I would be given an academic award based on how well I perform at school. _____

12. In order to get an academic award I would have to know the right people. _____

13. For me to be an outstanding student at school, it would take a lot of luck. _____

14. Getting recognized for my schoolwork would depend on how well I perform. _____

15. When required I can have a good deal of influence on my professor. _____

16. When I make plans at school, I am almost certain to make them work. _____

17. Although I might have the necessary abilities, I will not be given leadership responsibilities at school without appealing to those in positions of power. _____

18. It's not always wise for me to plan ahead at school because things turn out to be a matter of good or bad fortune. _____

19. When I get what I want at school, it's usually because I worked hard for it. _____

20. Whether or not I advance at school depends on whether I'm lucky enough to be in the right place at the right time. _____

Appendix F

Affect

Instructions. This questionnaire contains statements about your feelings toward those individuals whose performance you rated. Indicate the level you agree or disagree with each statement using the scale below:

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree

Rating

1. In general, I would like to spend more time with the people whose performance I rated. _____
2. In general, I tend to see eye to eye with the people whose performance I rated. _____
3. In general, I would like to get to know the people whose performance I rated better. _____
4. In general, I would regard the people whose performance I rated as good friends. _____

Appendix G

Self-efficacy

This section contains several statements about the rating process you participated in. Indicate the extent to which you agree or disagree with each statement, using the scale below:

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree

If a particular statement is not at all applicable to this organization, write NA in the space for ratings.

Rating

1. I understand those aspects of the rating process that I am responsible for participating in. _____
2. I have sufficient skills to execute adequately those steps of the rating process that I am responsible for. _____
3. I am sufficiently aware of the work behavior of individuals to evaluate them accurately. _____
4. I am not capable of rating the performance of individuals. _____

5. I am capable of using pertinent objective criteria to evaluate the quality of individuals' performance. _____

6. I believe I have the skills required to evaluate the quality of individuals' work. _____

7. I do not anticipate having difficulty evaluating the quality of individuals' work. _____

8. I do not seem capable of dealing with most problems that arise during the rating process. _____

9. When unexpected problems occur during the rating process, I do not handle them well. _____

10. I feel insecure about my ability to provide performance feedback to individuals. _____

11. When the rating process looks too complicated, I feel I want to give up on it. _____

Appendix H

Politics in Ratings

This section contains several statements about the rating process you participated in. These are general statements about how raters approach the rating process and do not necessarily reflect how you as a rater approach evaluation. Indicate the extent to which you agree or disagree with each statement, using the scale below:

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree

If a particular statement is not at all applicable, write NA in the space for ratings.

Ratings

1. Raters avoid giving ratings that may antagonize a ratee (e.g., a low rating). _____

2. Raters avoid giving low ratings because they fear that a ratee would want to leave their work group. _____

3. Raters inflate ratings of those people who are able to procure them special services or favors. _____

4. Raters inflate ratings of individuals who have access to valuable sources of information. _____

5. Ratings reflect in part a rater's personal liking or disliking of ratees. _____

6. Ratings are affected by the extent to which ratees are perceived as sharing the same values as the rater. _____

7. Ratings are affected by a ratee's ability to inspire enthusiasm in the rater. _____

8. Raters give ratings that will make themselves look good to their supervisors. _____

9. The rater-ratee personal relationship throughout the evaluation period (e.g., tense-relaxed; trusting-distrusting; friendly-hostile) affects ratings. _____

10. Raters avoid giving ratings that may have negative consequences for the ratee (e.g., no rewards). _____

11. Raters inflate ratings in order to maximize rewards for ratees (e.g.,rewards). _____

12. Raters produce accurate ratings only to the extent that they perceive that they may be rewarded for doing so or penalized for failing to do so. _____

13. Raters produce accurate ratings only to the extent that they perceive that this is the norm among other-raters. _____

14. Raters give low ratings to teach rebellious ratees a lesson. _____

15. Raters use ratings to send a message to their ratees (e.g., encourage risk taking, creativity, etc.). _____

16. Raters inflate ratings of those individuals who possess special characteristics. _____

17. Raters give higher performance ratings than are deserved in order to gain support or cooperation from their ratees. _____

Appendix I

Organizational Commitment

Instructions. This section contains several statements describing your feelings about this university. Indicate the extent to which you agree or disagree with each statement, using the scale below:

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Neither Agree nor Disagree	Slightly Agree	Moderately Agree	Strongly Agree

If a particular statement is not at all applicable to this organization, write NA in the space for ratings.

	Rating
1. The longer I stay at this university, the more I support its values and goals.	_____
2. For me, this is the best of all universities to be at.	_____
3. It would be difficult for me to adapt to a new university.	_____
4. I am proud to tell others that I am part of this university.	_____
5. Many problems would have to occur in order for me not to support the goals and values of this university.	_____

6. Often, I agree with this university's policies on important matters relating to its students. _____

7. Presently, it would be difficult for me if I decided to leave this university. _____

8. I talk up this university to my friends as a great university to be at. _____

9. I really care about the fate of this university. _____

10. I find my values and the university's values are very similar. _____

11. I am extremely glad I chose this university to attend over the others I was considering at the time I entered. _____

12. This university really inspires the best in me in the way of school performance. ____

13. I stick to this university because there is much to be gained. ____

14. In order to help this university be successful, I am willing to put in a great deal of effort beyond that normally expected from a student. ____

15. I am grateful for the opportunities this university has offered me. ____

Appendix J

Subject Information Sheet

Instructions. Please complete the following information as it pertains to you. Circle the appropriate response for each question.

Age

a) under 20 b) 20-25 c)25-30 d) over 30

Sex

a) Male b) Female

Indicate your year in school.

a) freshman b) sophomore c) junior d) senior e) other

Experience in Evaluating

Indicate the degree of experience you have in evaluating the performance of others and yourself using the scales below. A "1" on the scale means you have little experience evaluating performance, a "7" on the scale means you have a great amount of experience evaluating the performance of others or yourself.

The performance of others	1	2	3	4	5	6	7
Your own performance	1	2	3	4	5	6	7