

The Pennsylvania State University

The Graduate School

Department of Psychology

**IMPROVING REACTIONS TO PERFORMANCE APPRAISAL: THE EFFECTS
OF APPRAISAL SYSTEM CHARACTERISTICS, LEADER-MEMBER
EXCHANGE, AND TRUSTWORTHINESS**

A Dissertation in

Psychology

by

Anne Scaduto

© 2011 Anne Scaduto

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2011

The dissertation of Anne Scaduto was reviewed and approved* by the following:

Kevin Murphy
Professor of Psychology
Dissertation Adviser
Chair of Committee

Rick Jacobs
Professor of Psychology

Jeanette Cleveland
Professor of Psychology

Kelly Delaney-Klinger
Assistant Professor of Management

Melvin Mark
Professor of Psychology
Head of the Department of Psychology

*Signatures are on file in the Graduate School

ABSTRACT

The current study uses a sample of graduate student and adviser pairs to examine factors that shape reactions to performance appraisal. The purpose of this study was to determine how rater accountability, frequency of observation, justice perceptions, and trust in the appraisal system impact the relationship between LMX and performance ratings, and between trustworthiness and performance ratings. Another purpose was to determine how LMX and trustworthiness influence student reactions such as motivation to improve performance, accuracy ratings of the performance appraisal, and satisfaction of performance ratings. Results of the study suggest a relationship between LMX, trustworthiness, performance appraisal system characteristics, and overall performance ratings. Additionally, LMX ratings appeared to become integrated with ratings of overall performance, which in turn influence students' ultimate feedback reactions to performance appraisals. By focusing on improving performance appraisal system characteristics, organizations have the opportunity to influence their employees reactions to feedback and improve overall organizational effectiveness. Limitation, future research and practical implications are discussed.

TABLE OF CONTENTS

List of Figures	vi
List of Tables	vii
Acknowledgements	viii
Chapter 1 Introduction	1
Model	3
Chapter 2 Literature Review	8
Why Employees React to Performance Appraisal	8
Performance Appraisal System Characteristics	10
Organizational Justice	10
Rater Accountability	16
Frequency of Observation	18
Leader-Member Exchange	20
Trustworthiness and Leadership	23
Trust in the Appraisal System	26
Purpose of Current Research	28
Chapter 3 Methodology	30
Participants	30
Measures	31
Procedure	36
Chapter 4 Results	37
Reliability of Scales	38
Confirmatory Factor Analysis	40
Unit of Analysis	41
Statistical Analysis	47
Methods of Mediation	54
Chapter 5 Discussion	62
Limitations and Future Reserach	67
Practical Implications and Conculsions	69

BIBLIOGRAPHY	72
Appendix A Student Survey	80
Appendix B Adviser Survey	83
Appendix C Adviser Recruitment Letter	87
Appendix D Student Recruitment Letter	88
Appendix E Moderation Analyses: Student Ratings	89
Appendix F Moderation Analyses: Adviser Ratings	94
Appendix G Mediation Analyses: Student Ratings	99
Appendix H Mediation Analyses: Adviser Ratings.....	105

LIST OF FIGURES

Figure 1. High Level Hypothesized Model..... 4

Figure 2. Specific Variable Hypothesized Model..... 5

LIST OF TABLES

Table 1: Means, Standard Deviations, Reliabilities.....	39
Table 2: Correlations between Performance Appraisal System Characteristics, Leader-Member Exchange, Trustworthiness, Performance Ratings and Feedback Reactions (Individual Level).....	45
Table 3: Correlations between Performance Appraisal System Characteristics, Leader-Member Exchange, Trustworthiness, Performance Ratings and Feedback Reactions (Dyad Level).....	48
Table 4: Moderated Multiple Regression: OCBs as the Dependent Variable.	49
Table 5: Moderated Multiple Regression: Timeliness as the Dependent Variable	50
Table 6: Moderated Multiple Regression: Quality as the Dependent Variable	51
Table 7: Moderated Multiple Regression: Dependability as the Dependent Variable.....	52
Table 8: Moderated Multiple Regression Overall Performance as the Dependent Variable.....	53
Table 9: Performance Ratings Mediating the Relationship between LMX and Satisfaction	59
Table 10: Performance Ratings Mediating the Relationship between LMX and Motivation.....	59
Table 11: Performance Ratings Mediating the Relationship between Trustworthiness and Satisfaction	60
Table 12: Performance Ratings Mediating the Relationship between Trustworthiness and Motivation	60

ACKNOWLEDGEMENTS

This dissertation would not have been completed if it were not for the help of so many people. I would like to thank my chair, Kevin Murphy, for his guidance, patience, and support throughout my graduate career. I would also like to express my gratitude to my committee members; Dr. Jeanette Cleveland, Dr. Rick Jacobs, and Dr. Kelly Delaney-Klinger for their valuable opinions and comments. Most importantly, I would like to thank my parents for encouraging me my whole life.

Chapter 1

Introduction

Performance appraisal is a formal process used to evaluate what individuals do on the job (Brannick & Levine, 2002). Performance appraisal occurs in organizations for a number of purposes including administrative decisions (i.e. promotion, raise), and feedback and development (Murphy & Cleveland, 1995). Therefore, the performance evaluation process is central to many personnel decisions (Landy, Barnes, & Murphy, 1978). Because of the pivotal role performance appraisal plays in human resource management, there has been a significant amount research conducted in order to understand appraisal systems (Mayer & Davis, 1999).

The following study explores variables that are believed to play a critical role in the performance appraisal system based on past research such as leader-member exchange, trustworthiness, and performance appraisal system characteristics including frequency of observation, justice perceptions, trust in system, and rater accountability. More specifically, this research is interested in discovering which of these variables impact both self and leader ratings of performance and reactions to these ratings such as motivation to improve performance, perceptions of accuracy, and satisfaction. These specific variables were chosen because there has been much research theorizing relationships between them. However, although there has been some empirical performance appraisal research studies conducted using these variables; no study has put them all in one model. Moreover, empirical studies which include feedback reactions are

minimal. Due to the fact that performance appraisal is such a prevalent process with many implications, it is important to understand all of its components. Furthermore as discussed in more detail below, performance appraisal research often ignores these types of variables which may prove to be quite important in terms of overall organizational effectiveness.

Scientists and practitioners have historically been concerned with issues in performance appraisal such as appraisal effectiveness and the measurement of performance (Landy & Farr, 1980). Consequently, much of the performance appraisal research in industrial and organizational psychology has focused on rater errors and accuracy of ratings. However, the message has been increasingly clear that performance appraisal takes place in a social context which plays a major role in the effectiveness of such appraisals and in participants' reactions to the performance appraisal process (Bretz, Milkovich, & Read, 1992). "Performance appraisals are no longer just about accuracy, but are about much more including development, ownership, input, perceptions of being valued, and being part of an organizational team" (Levy & Williams, 2004, p. 889). Moreover, there has not been much emphasis on employee reactions to performance appraisal. In fact, Murphy and Cleveland (1995) referred to reaction criteria as one class of "neglected criteria" in performance appraisal system evaluation and research (p. 310). However, both academic and practitioner reports suggest that many employees are considerably dissatisfied with the appraisal process in their organization (Elicker, Levy, & Hall, 2006). According to Keeping and Levy (2000) satisfaction has been the most frequently measured appraisal reaction and is most often conceptualized in three ways: 1) satisfaction with the appraisal session, 2) satisfaction with the appraisal system, and 3)

satisfaction with the performance ratings. Employee reactions found in the literature other than satisfaction include reactions such as perceived fairness and accuracy, often operationalized in terms of organizational justice, and motivation to improve performance (Cawley, Keeping, & Levy, 1998). Importantly, Jacobs, Kafry, and Zedeck (1980) identified performance appraisal reactions as a criterion of performance appraisal effectiveness.

Notwithstanding the complexity that the social context adds to performance appraisals, there is also a major gap between science and practice in this area. According to Thomas and Bretz (1994), there are two fundamentally different perspectives on performance appraisal which have created this gap between research and practice. While managers often focus on the processes and behaviors of performance appraisal such as its fairness and usefulness, researchers focus more strongly on cognitive aspects of the rating process. Thomas and Bretz (1994) report a disappointing result from a survey of Fortune 100 firms revealing that despite identifying issues related to fairness in appraisal systems, only one-third of the organizations surveyed actually conduct attitude surveys to assess these types of variables. The current study attempts to bridge this gap by emphasizing not only the importance of feedback reactions in regards to performance, but what particular factors shape these reactions.

Model

Figure 1 below is the overall model of the current study, and Figure 2 breaks down the constructs into their components. The major constructs include Leader-member exchange (contribution, loyalty, affect, professional respect), trustworthiness (ability, benevolence, integrity), performance appraisal system characteristics (justice, frequency

of observation, rater accountability, trust in the appraisal system), performance ratings (leader ratings, self ratings), and feedback reactions (motivation to improve performance, satisfaction, accuracy).

The model tested in this study asserts that there is a positive relationship between leader-member exchange (LMX) or the dyadic relationship between supervisor and subordinate and trustworthiness. More specifically, the LMX component *loyalty* is expected to be positively related to the trustworthiness component *benevolence*, and the LMX component *professional respect* is expected to be positively related to the trustworthiness component *ability*.

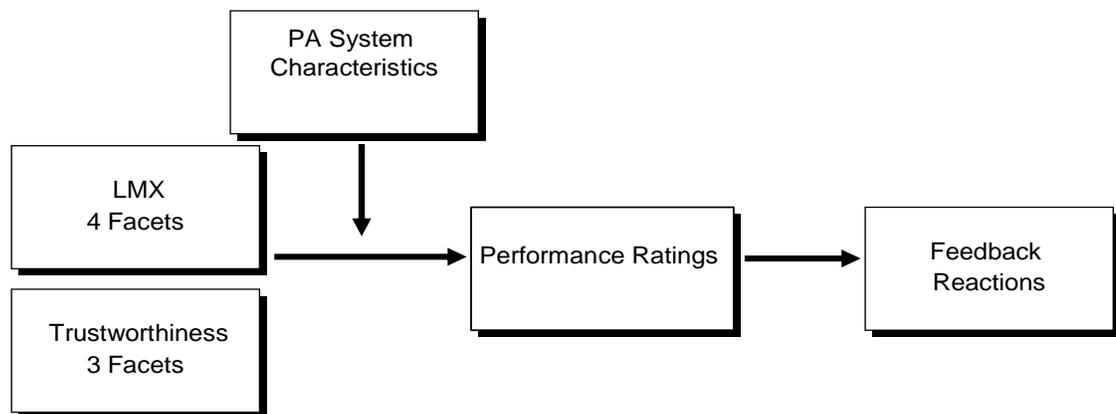


Figure 1. High Level Hypothesized Model

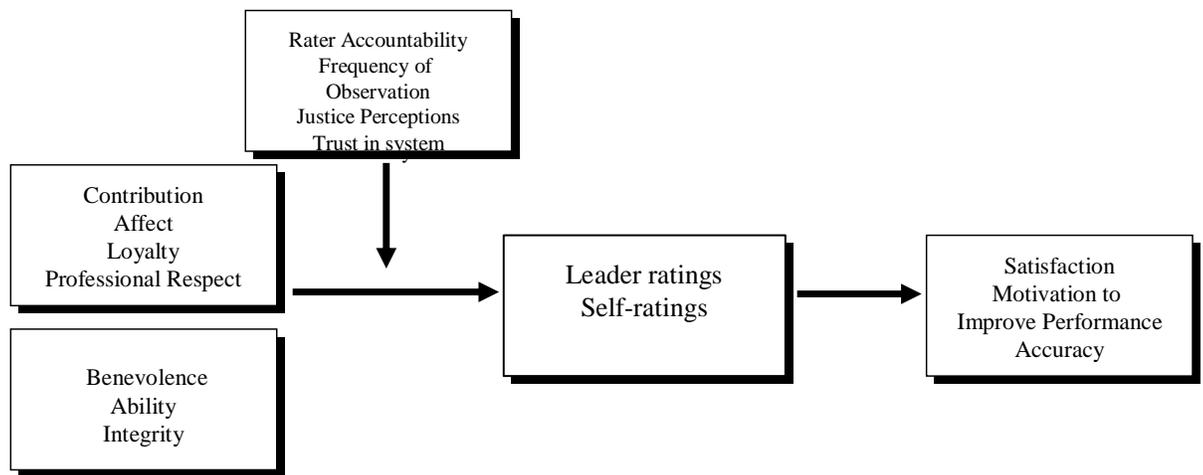


Figure 2. Specific Variable Hypothesized Model

It is believed that rater accountability, or the expectation that the rater may be called on to justify one's beliefs, feelings, and actions to others, (Lerner & Tetlock, 1999) is positively related to frequency of employee observation, or the number of times raters spend observing rates. If raters are concerned that they will be held accountable for their performance ratings, then frequency of observation becomes an important factor for justifying the ratings. The more accountable a rater is held the more likely he or she will ensure that the ratee is being observed as frequently as possible in order to provide

sufficient evidence of performance ratings. In this study frequency of observation serves as a proxy for whether or not the rater has a solid basis for his or her appraisal rating. Justice perceptions, specifically procedural justice, and trust in appraisal system are believed to be related to rater accountability and frequency of observation as well.

This model additionally proposes that performance appraisal system characteristics are positively related to LMX. For example, increased levels of frequency of observation lead to a stronger relationship between supervisor and subordinate in which there is more communication, increased resources, advantageous treatment from the supervisor, and ultimately higher performance.

The model predicts that performance appraisal system characteristics moderates the relationship between LMX and performance appraisal ratings such that higher levels of rater accountability, frequency of observation, justice perceptions, and trust strengthen the relationship between LMX and performance ratings and low levels of these variables weaken the relationship. It also predicts that performance appraisal system characteristics moderate the relationship between trustworthiness and performance appraisal ratings such that higher levels of rater accountability, frequency of observation, justice perceptions, and trust strengthens the relationship between trustworthiness and performance ratings and low levels of these variables weakens the relationship. Additionally, performance ratings from both the adviser and student partially mediate the relationship between LMX and feedback reactions and trustworthiness and feedback reactions.

These key variables are linked in the model in this way based on previous studies and theoretical conceptualizations. The following sections review the literature that supports the development of this model.

Chapter 2

Literature Review

Why Employees React to Performance Appraisal

More recently performance appraisal research has begun to explore the effects of the social context on the performance appraisal process. Levy and Williams (2004) reviewed over 300 articles and developed a model which broadens the conceptualization of performance appraisal effectiveness to include and strongly highlight employee reactions to feedback and the feedback environment; variables such as managers and employees feeling comfortable both providing and receiving feedback. Researchers have argued that the organization's feedback environment should play a vital role in how feedback is sought, perceived, processed, accepted, used, and reacted to. The current research takes on this same view as it examines similar variables. According to the Levy and Williams (2004) review, the feedback environment can be measured by looking at the employee's perceptions of feedback source credibility, feedback quality, feedback delivery, frequency of feedback, source availability, and the extent to which feedback seeking is encouraged (Levy & Williams, 2004). These authors conclude that researchers are now more aware of the importance of the social context within which performance appraisals operate. The current study takes some of these theoretical arguments to an empirical level by measuring variables that are similar to the components of the feedback

environment, such as performance appraisal system characteristics, LMX, and trustworthiness, and ultimately reactions to feedback.

Some of the variables described above have been hypothesized to be related to feedback reactions. Additionally, negative reactions to performance appraisals have led researchers to search for explanations as to why employees react so differently to evaluation. Classic explanations hypothesize and intuitively suggest that employees will react more positively when they receive higher ratings of performance, or when ratings are higher than originally expected (Elicker et al., 2006). However, this would imply that all employees receiving the same feedback would have similar reactions. Because this is unlikely due to individual differences and other confounding variables, researchers believe there must be alternate factors related to positive and negative reactions to performance appraisals, regardless of the nature of the feedback and the current study tested this belief.

Moreover, we do know that performance appraisal has broad implications for attitudes and behaviors in organizations. Erdogan's (2002) theoretical piece suggests that antecedents of employee justice perceptions of performance appraisal systems include due process characteristics, organizational culture, pre-appraisal leader-member exchange (LMX), perceived organizational support (POS), impression management behaviors of raters, perceived basis of LMX, and perceived type of information raters use. The author used social exchange and accountability theories to link justice perceptions to organizational, leader-related, and performance-related outcomes. Erdogan's (2002) proposed model identifies directions for future research as well as the current study. Many of the variables focused on in this theoretical piece map directly onto

the model in Erdogan's (2002) piece. The following sections discuss research around performance system characteristics, trustworthiness, LMX, and feedback reactions in detail.

Performance Appraisal System Characteristics

The current study uses the term Performance Appraisal System Characteristics to refer to organizational justice (procedural and distributive), rater accountability, frequency of observation, and trust in the performance appraisal system. Trust in the performance appraisal system will be discussed in a later section.

Organizational Justice

Organizational justice is a term used to describe the role of fairness as it is related to the workplace (Moorman, 1991). More specifically, organizational justice deals with ways in which employees determine if they have been treated fairly and the ways in which those perceptions influence other work-related variables such as organizational citizenship behaviors (Moorman, 1991).

In the organizational justice literature, three forms of justice have been identified (Colquitt, Conlon, Wesslen, and Porter, 2001). Distributive justice refers to fairness of the outcomes an individual receives. Often it is looked at as inputs versus outputs. For example, employees may assess the fairness of their appraisals by comparing the relative ratings they received in return for the work they contributed. Procedural justice refers to

the fairness of procedures used to arrive at a decision. Lastly, interactional justice refers to the perceived fairness of interpersonal communication related to procedures.

Interactional justice consists of two types of justice: interpersonal and informational (Colquitt et al., 2001). Interpersonal justice is linked to the treatment of people. For example, being polite and respectful when executing procedures or determining outcomes. Informational justice refers to the explanation of why procedures were used, or why outcomes were distributed in a particular way (Colquitt et al., 2001).

Because performance evaluations can be subjective in many jobs, researchers have realized the greater potential for perceptions of unfairness. For example, during performance appraisals, raters are cognitively limited and there is discrepancy on what constitutes good performance (Folger, Konovsky, & Cropanzano, 1992). Performance appraisal researchers are realizing the importance of measuring ratee justice perceptions to assess how successful a performance appraisal system really is; therefore serving as criteria of effectiveness in performance appraisals (Erdogan, Kraimer, & Liden, 2001). Erdogan et al. (2001) provided support for the construct validity of a two dimensional conceptualization of procedural justice; system procedural justice and rater procedural justice. This study demonstrated that employees make source attributions when making conclusions about procedural justice. According to these authors, it is possible that these attributions may provide a link between employee perceptions of justice and future employee behaviors (Erdogan et al., 2001).

Measuring justice perceptions is additionally important because they have been linked to important organizational outcomes. For example, perceptions of procedural justice have been found to be related to organizational citizenship behaviors, employee

attitudes, and retaliation behaviors (Erdogan et al., 2001). More specifically, justice perceptions have been linked to attitudes such as organizational commitment (Folger & Konovsky, 1989; Konovsky & Cropanzano, 1991), trust in management and turnover intentions (Konovsky & Cropanzano, 1991), and behaviors such as employee performance (Konovsky & Cropanzano, 1991), altruism, courtesy, sportsmanship and conscientiousness (Moorman, 1991).

Landy et al. (1978) found that giving the ratee an opportunity to express his or her feelings during performance appraisals predicted a measure of perceived fairness and accuracy of the evaluation. The questionnaire used in this study to measure fairness and accuracy was composed of items dealing with areas such as the effect of criticism and praise, salary discussion based on performance appraisal results, and reasons for performance appraisal (Landy et al., 1978). The results of this study indicate that perceptions of fairness and accuracy of performance appraisal are significantly related to process variables. Moreover, fair and accurate appraisals occurred when supervisors evaluated performance often, were familiar with the performance levels of the ratee, agreed with the ratee on job duties, and aided in ratee development (Landy et al., 1978). Therefore, it seems clear that the way in which the performance appraisal is carried out in terms of both overall system procedures and rater procedures, affects the attitudes and perceptions of the ratee. According to Lawler's (1967) model of variables that might affect the validity of performance ratings, the "objective characteristics of the evaluation system, individual differences, and organization characteristics combine and filter through these attitudes that vary in validity" (Landy et al., 1978, p. 753). More relevant to the current study, Lawler (1967) discusses a variable labeled "attitudes toward fairness

and acceptability of the system.” The Landy et al., (1978) study helped operationalize this variable further demonstrating ways to obtain positive reactions from employees about the performance appraisal system.

Some have suggested that justice perceptions may have an effect on performance; specifically organizational citizenship behaviors. Organizational citizenship behaviors (OCBs) are optional behaviors that help other members of the organization perform their job, or show support for and conscientiousness toward the organization (Borman & Motowidlo, 1993). According to Organ (1988), there are five OCB dimensions: altruism, courtesy, sportsmanship, conscientiousness, and civic virtue. Moorman (1991) looked at the relationship between perceptions of fairness and organizational citizenship behaviors in two organizational samples. The results of this study found support for a relationship between fairness perceptions (particularly those from interactional justice) and four of the five OCB dimensions (Moorman, 1991). The strongest implication of Moorman’s (1991) study is that supervisors can directly impact employees’ citizenship behaviors. More specifically, this research suggests that if managers want to increase citizenship behaviors, they should work to increase the fairness of their interactions with employees, such as when conducting performance appraisals. More evidence for the perceptions of fairness-OCB relationship comes from a meta-analysis conducted by Organ and Ryan (1995). The meta-analysis revealed that job attitudes such as perceived fairness are robust predictors of OCBs. Because most of the studies used in this meta-analysis attempted to measure distributive and procedural fairness separately and only some measured interactional justice, Organ and Ryan (1995) opted to look at the relationship between overall fairness and OCBs.

Research in the past decade has been extremely interested in organizational justice, which has led to an expansion of the issue of fairness. In the context of performance appraisal, Greenberg (1986) defines distributive justice as concerned with the fairness of performance ratings relative to work performed. This is in contrast to procedural justice which involves the perceptions of the appraisal process itself. Put a different way, distributive justice refers to the perceived fairness of the amounts of compensation employees receive, while procedural justice refers to the perceived fairness of the means used to determine those amounts (Folger & Konovsky, 1989). In Greenberg's study (1986), middle managers from three organizations responded to an open-ended questionnaire that asked them to describe the determinants of fair or unfair performance appraisals. Greenberg (1986) grouped the responses from the questionnaire and identified five categories of procedural fairness (rater familiarity with ratee's work, consistent application of standards, seeking input before an evaluation and using it, two-way communication during an interview, and the ability to contest an evaluation) and two categories of distributive justice (receipt of rating based on performance achieved and recommendation for salary/promotion based on rating). Participants rated the importance of each of these determinants on a nine-point scale. Both distributive determinants were rated as more important than all of the procedural ones. However, based on these responses it is clear that both aspects of organizational justice need to be taken into account when conducting performance appraisals. A limitation to this study is that all three organizational samples were very similar.

Subsequently, Folger and Konovsky (1989) looked at the impact of distributive and procedural justice on the reactions of employees to decisions about pay raises. These

researchers found distributive justice to account for more unique variance in satisfaction with pay than did procedural justice. However, procedural justice accounted for more unique variance in trust in supervisor. Additionally, Korsgaard and Roberson (1995) discovered that fair procedures can create positive attitudes toward decisions that might be otherwise viewed negatively. For example, perceptions of voice, or the level of employee participation in decision making, and distributive justice were related to trust in manager (Korsgaard and Roberson, 1995). In addition, Folger et al. (1992) discussed a due process metaphor, similar to dispute resolution, to help alleviate some problems associated with performance appraisal resulting in an increased fairness as seen by employees. These authors suggest that implementing performance appraisal interviews would allow discussion and also provide an opportunity for employees to bring evidence to support their performance levels in relation to others in hopes of improved evaluations (Folger et al., 1992). In summary, many researchers and many studies are focusing on the importance of organizational justice as a predictor of OCBs, employee reactions, and overall perceptions of fairness. Relationships such as these within the context of performance appraisal systems are the focus of the current study.

This study focused on how justice perceptions relate to other appraisal system characteristics, leader-member exchange, trustworthiness, performance ratings, and feedback reactions.

Rater Accountability

The performance appraisal review by Levy and Williams (2004) demonstrates that proximal process variables such as accountability have a direct impact on how the appraisal process is conducted. In the context of performance appraisal, accountability is typically thought of as the extent to which a rater is held answerable to someone else for his or her ratings of another employee (Levy & Williams, 2004). Accountability refers to the implicit or explicit expectation that the rater may be called on to justify one's beliefs, feelings, and actions to others (Lerner & Tetlock, 1999). Accountability usually implies that individuals who do not provide an adequate justification for their actions will suffer negative consequences (Lerner & Tetlock, 1999). Conversely, individuals who do provide convincing justifications will experience positive consequences (Lerner & Tetlock, 1999).

According to Curtis, Harvey, and Ravden (2005) the operationalization of accountability varies widely in the accountability literature. There are two major forms of accountability: upward accountability and downward accountability. Much of the accountability literature focuses on downward accountability, or being held accountable to the ratee. This is often defined as having to give feedback and/or justify performance appraisal ratings to the subordinate (Curtis et al., 2005). Downward accountability is quite common in organizations. For example, Mero and Motowidlo (1995) found that it is often the case that a feedback meeting between the supervisor and the subordinate is required after the performance appraisal is completed. Although this is often policy, the extent to which these meetings occur is unknown (Mero & Motowidlo, 1995).

While downward accountability has been found to increase leniency error or the tendency to rate employees too positively, upward accountability will serve to decrease leniency error (Curtis et al., 2005). Upward accountability, or raters being held accountable for their ratings of their subordinates to a superior, is not quite as common in organizations, or researched very often. When raters believe that a superior will carefully review ratings of their subordinate, their motivation to be accurate should potentially increase because of their desire to appear competent (Harris, Smith, & Champagne, 1995). The exception is when raters perceive that the person to whom they are accountable expects certain ratings (Tetlock, 1985). More evidence for reduced leniency in ratings comes from the Curtis et al. (2005) study in which holding raters accountable lead to a reduction in the leniency of ratings; particularly when the purpose of the appraisal was administrative in nature.

If raters are only held accountable to the ratee, as with downward accountability, they should be heavily influenced by the thought of having to explain the ratings to the ratee face-to-face (Klimoski & Inks, 1990). Research has demonstrated that the desire to avoid the tension of having to give negative feedback overrides a rater's desire to be accurate (Tesser & Rosen, 1975). On the other hand, if raters are held accountable to their superiors, in the case of upward accountability, they should want to rate more accurately to appear competent (Simonson & Nye, 1992), and the desire to avoid tension associated with the ratee should not be present. Interestingly, the result of holding raters upwardly and downwardly accountable simultaneously is not yet known. Curtis et al. (2005) hypothesizes that if raters are held accountable to the ratee and to their superiors, they should feel torn. On one hand, they should want to appear competent and help the

organization. On the other hand, however, they should want to avoid giving negative feedback to the ratee. Ratees should feel forced to choose between the two options or choose a middle ground (Curtis et al., 2005).

It is also important to examine what factors influence accountability. In a study conducted by Rutkowski and Steelman (2005), factors influencing accountability such as LMX were examined because accountability has recently been a focus in the feedback literature. However, the context of this particular study is quite different from the proposed study. For example, these authors looked at how managers used feedback from subordinates to improve their future performance. The goal of the upward feedback process is to present managers with feedback in order to improve their leadership behaviors. It was hypothesized that LMX would be an antecedent to this feeling of accountability and the results of the study found this to be the case.

The current study focused on both downward and upward accountability, and how this variable related to other appraisal characteristics, leader-member exchange, trustworthiness, performance ratings, and feedback reactions.

Frequency of Observation

In 1967, Lawler proposed a model of the variables that might affect the validity of performance ratings. In that model, the variable “attitudes toward fairness and acceptability toward the system” played a major role (Lawler, 1967). In an empirical study conducted by Landy et al. (1978), these researchers helped to operationalize this important parameter from Lawler’s model. In their study, employee perceptions of

fairness and accuracy of a performance appraisal system were examined via a questionnaire given to managerial and professional employees from a large manufacturing organization. The results of this study indicated that a performance evaluation is considered both fair and accurate when supervisors evaluate performance frequently, are familiar with the performance level of the individual being evaluated, are in agreement with the subordinate on job duties, and engage in helping the subordinate form plans for eliminating performance weaknesses (Landy et al., 1978).

The results of this study demonstrate strong support for the Lawler (1967) model. It appears that the way in which an evaluation is carried out affects the attitudes of the individual being evaluated (Landy et al., 1978). More specifically, the results confirm some more traditional ideas about performance appraisals. For example, relevant to the current study, the Landy et al. (1978) research indicates that supervisors should prepare carefully for the appraisal process. Supervisors can prepare by agreeing with the subordinate on responsibilities and, more relevant to the current study, by devoting sufficient amounts of time to observe the subordinate's performance (frequency of observation). It seems logical to state that frequency of observation is an indicator of whether or not the rater has a sound reason for his or her appraisal ratings. Although there is no doubt that these variables are related to each other, a longitudinal study with both pre- and post- assessments would be necessary in order to establish if perceptions of fairness and accuracy follow more frequent observations. (Landy et al., 1987).

Intuitively, the relationship between frequency of observation and rater accountability should be a positive one. If raters are concerned that they will be held accountable for their ratings to subordinates, then frequency of observation becomes an

important variable for the justification of the ratings. The more accountable a rater is, the more likely he or she will ensure that the ratee is being observed as frequently as possible. This relationship between frequency of observation and rater accountability was tested in the current study. Additionally, the relationship between frequency of observation, other appraisal system characteristics, trustworthiness, performance ratings, and feedback reactions was also tested.

Leader-Member Exchange

Leader-member exchange (LMX) theory was developed to capture the process through which leaders respond to and interact with subordinates (Levy & Williams, 2004). The theory suggests that leaders interact, respond to, and treat subordinates differentially depending on the subordinate's membership in the "in" or "out" group (Dansereau, Graen, & Haga, 1975). Dienesch and Liden (1986) and Liden and Maslyn (1998) proposed four dimensions of LMX relationships termed *contribution* (e.g., performing work outside of what is depicted in the job description), *affect* (e.g., friendship and liking), *loyalty* (e.g., loyalty and mutual obligation), and *professional respect* (e.g., respect for professional abilities). Additional LMX research has produced measures of these constructs and demonstrated validity of these dimensions (Liden & Maslyn, 1998; Schriesheim, Neider, Scandura, & Tepper, 1992). LMX relationships are likely composed of one or more of these factors; therefore examining the different dimensions of LMX is important.

LMX theory (Dansereau et al., 1975) further describes the supervisor-subordinate relationship as a dyadic social exchange process that is unique to each supervisor-subordinate pair (Graen & Uhl-Bien, 1995). This theory suggests that a subordinate in a high-LMX dyad will receive more attention and resources from his or her supervisor, and will likely respond to this advantageous treatment with higher performance, trust, loyalty, and positive attitudes (Elicker et al., 2006). These relationships tend to be determined early and to be somewhat stable (Liden, Wayne, & Stilwell, 1993). Paralleling this more rewarding working relationship in high-LMX supervisor-subordinate pairs, there is also a higher degree of mutual liking than is seen with low-LMX relationships (Engle & Lord, 1997). An important point to highlight is that higher quality supervisor-subordinate relationships exhibit more open and friendly communication (Elicker et al., 2006). In addition, leader-member exchange theory hypothesizes that the quality of the relationship developed between a leader and a follower is predictive of outcomes at the individual, group, and organizational levels (Graen & Schieman, 1978).

In their performance appraisal review, Levy and Williams (2004) demonstrate that proximal process variables such as LMX have a direct impact on how the appraisal process is conducted. Further, research conducted by Duarte and colleagues (Duarte & Goodson, 1994; Duarte et al., 1993) demonstrated that the relationship between objective measures of performance and supervisor ratings of performance are moderated by LMX such that in-group members were rated higher despite objective levels of performance. Across these two studies, these authors found that individuals who communicated with their supervisor more frequently and were in a high LMX relationship received the highest performance ratings. An interesting finding related to this is that individuals with

high communication frequency and in low LMX relationships received the lowest performance ratings (Duarte & Goodson, 1994; Duarte et al., 1993).

In an attempt to discover and test factors influencing employee reactions to performance appraisals, Elicker et al. (2006) developed a model of feedback reactions that emphasized the importance of the relationship between a supervisor and a subordinate in determining employee reactions. The results demonstrated that the effect of a good supervisor-subordinate relationship results in differing behaviors and more positive cognitive evaluations of performance appraisals (Elicker et al., 2006). Therefore, these results suggest that employee reactions appear to differ based on differences in perceived treatment. Positive employee reactions to performance appraisals in high-quality exchange relationships seem to greatly depend on their treatment in the appraisal session and the resulting justice judgments (Elicker et al., 2006). More specifically, the results of the Elicker et al. (2006) study demonstrate the importance of exchange quality as it is related to feedback session reactions such as motivation to improve, and perceptions of accuracy and utility. This study indicates that justice judgments have implications not only for reactions to performance appraisals, but broader outcomes as well.

In regards to the model, frequency of observation was hypothesized to be positively related to LMX. An increased number of observations are equivalent to more time spent together. The more time supervisors and subordinates spend together, the stronger the dyadic social exchange will likely become. Additionally the model hypothesized a positive relationship between downward accountability and LMX. LMX

components were also predicted to be related to trustworthiness, justice perceptions, performance ratings, and feedback reactions.

Trustworthiness and Leadership

Researchers have hypothesized that trust is a key component in managing the supervisor-subordinate relationship. There is agreement that without trust, social exchange relationships cannot be developed. As a result, researchers have used trust levels as an indicator of the existence of social exchange relationships (Colquitt et al., 2007). Additionally, the trustworthiness antecedents may encourage a social exchange relationship (Colquitt et al., 2007). Colquitt et al.'s (2007) meta-analysis summarized the relationship between trust variables and job performance. The results of the Colquitt et al. (2007) study indicated moderately strong relationships between trust and performance outcomes. For example, trust was positively related to task performance and organizational citizenship behaviors, but negatively related to counterproductive work behaviors (Colquitt et al., 2007).

Trust that individuals have in their leaders is an important concept in applied psychology (Dirks & Ferrin, 2002). In the context of performance appraisals, if an employee believes a supervisor has the skills to appraise properly, has the interest of the employee at heart, and believes the supervisor upholds standards and values, the employee is likely to trust that supervisor. It has been argued by researchers (Argyris, 1964) that trust for management is tied to important productivity-related outcomes which are considered a critical component of task performance. Consequently, Mayer and Davis

(1995) proposed that “the effective use of performance appraisal systems may provide an opportunity to build trust in organizations” (p. 123).

Dirks and Ferrin (2002) examined the findings and implications of the research on trust in leadership from the past four decades. Through their research it was discovered that direct leaders, such as supervisors, appear to be an important referent of trust (Dirks & Ferrin, 2002). Specifically, trust in one’s direct leader had strong, positive correlations with work attitudes and intention to remain with the organization. Trust was also moderately positively correlated with task and citizenship-related performance. Maybe more importantly for the context of performance appraisal, trust had positive correlations with distributive and procedural justice, and participation in decision making (Dirks & Ferrin, 2002).

Continuing with the idea of trust contributing to ratee perceptions of evaluations, Mayer and Davis (1995) conducted a quasi-experiment looking at trustworthiness and perceptions of a performance appraisal system. The authors proposed three factors of trustworthiness (ability, benevolence, and integrity) based on Mayer et al.’s (1995) study. Ability is defined as a group of skills, competencies, and characteristics that allow an individual to have influence within a specific domain. In the context of an organization, ability is the level of formal and informal influence managers or supervisors are perceived to have, and also their perceived competence. Benevolence is the extent to which a trustee is believed to want to do good to the trustor without a selfish motive. An example of benevolence is if an employee believes his or her manager cares about his or her interests. Integrity is defined as the trustor’s perception that the trustee follows a set of principles consistently and that the trustor finds this particular set of principles

acceptable. It was found that all three of these factors mediated the relationship between perceptions of the appraisal system and trust and intuitively so (Mayer and Davis, 1995).

There has also been an attempt to describe the relationship between trust and leadership styles with theories such as transformational and transactional leadership. Leadership styles such as these are also tied to LMX theory. For example, transformational leaders are better able to articulate the importance and values associated with preferred outcomes in ways that are easily understood, while also conveying higher levels of expectations for the followers (Avolio & Bass, 1988). Many leadership researchers have emphasized the development of a shared vision as being an essential part of the transformational leadership process (Jung & Avolio, 2000). During this shared vision development, leaders are seen as role models and motivators, and followers typically demonstrate a higher degree of trust in them (Gardner and Avolio, 1998). In fact, a follower's trust in the leader is one of the most important variables mediating the effectiveness of this type of leadership (Yukl, 1998). Based on this description, relationship-based trust appears to map onto transformational leadership quite well.

In contrast, transactional leadership is described as a simple exchange process between leaders and followers. A transactional leader recognizes a follower's expectations and provides rewards in exchange for that follower's performance (Jung & Avolio, 2000). It is important to note here that unlike in transformational leadership; in transactional leadership it is unnecessary for followers to develop a sense of trust in their leader. Although transactional leadership has been shown to be effective, a transactional leader is not looking to personally develop his or her follower (Jung & Avolio, 2000).

The results of the Jung and Avolio (2000) study demonstrate that transformational leadership has both direct and indirect effects on followers' performance, while transactional leadership only has indirect effects on performance. Followers' trust had a mediating effect on the relationship between transformational leadership and various measures of performance. More specifically, transformational leadership had a strong positive effect on performance quality (Jung & Avolio, 2000). In addition, the relationship between transformational leadership and various measures of performance was mediated by followers' trust (Jung & Avolio, 2000). These findings have obvious implications for performance appraisal systems. For example, it seems that trusting the appraiser may lead to more positive perceptions of the performance appraisal process in general in addition to actual increased performance on the job.

This study tested the relationship between trustworthiness, its components, and LMX, appraisal system characteristics, performance ratings, and feedback reactions.

Trust in the Appraisal System

The last performance appraisal system characteristic is *trust in the appraisal system*. In conjunction with the idea that trust in leader is an important factor in performance appraisal system effectiveness, often defined as positive perceptions, Farr and Jacobs (2006) states that "perhaps the construct with the greatest impact on the effectiveness of a performance appraisal system in an organization is what we might label as the "Collective Trust of Stakeholders in the Appraisal Process"(p. 5-6). In other words, many individuals must *trust* the process before performance appraisal can be effectively

implemented in an organization. Farr and Jacobs (2006) say that “trust is a central and critical factor in the effective development and implementation of performance appraisal systems in organizations” (p. 6). The degree to which the rater is seen as honest and high on integrity may change the perceived effectiveness of the performance appraisal system. The rater is an essential part contributing to a successful appraisal process (Farr & Jacobs, 2006). Not only does the rater have organizational needs that need to be met, but he or she also has personal goals regarding trust and the need to be seen as trustworthy by ratees (Farr & Jacobs, 2006).

According to Farr and Jacobs (2006), trust in the rater can come in the form of a) knowing that the performance appraisal process will be administered consistently across ratees, b) ensuring that the information produced will be used in the manner communicated to all involved, c), making sure that appeals and new information are heard correctly, d) creating an environment where communication about results is well understood. A hypothesis in this study is that increasing trust in those that are being evaluated will lead to positive outcomes in the appraisal process (Farr & Jacobs, 2006).

Dirks and Ferrin (2001) conducted a meta-analysis on the effects of interpersonal trust on workplace-related attitudes. These researchers discovered that trust has both direct and moderating effects on variables related to appraisal systems (Dirks & Ferrin, 2001). For example, it was found that higher levels of interpersonal trust were associated with higher levels of perceived accuracy and fairness of performance appraisals (Dirks & Ferrin, 2001). In addition to this direct effect, trust moderated the relationship between performance feedback and later individual performance (Dirks & Ferrin, 2001). The combination of all of these findings suggests that greater trust may aid in effective

implementation of a performance appraisal system (Farr & Jacobs, 2006). In the concluding comments section of the Farr and Jacobs (2006) chapter, the authors state that “clear communication and trust in the process [of performance appraisal] must be met for the appraisal process to influence future work behavior and attitudes at the individual, team, and enterprise level” (p. 26). Trust in rater is a central component of the overall idea of trust in the performance appraisal process. The current study hypothesized a relationship between trust in the appraisal system and LMX, trustworthiness, other appraisal characteristics, performance ratings, and feedback reactions.

Purpose of Current Research

The purpose of the current study was to determine how rater accountability, frequency of observation, justice perceptions, and trust in the appraisal system impact the relationship between LMX and performance ratings, and between trustworthiness and performance ratings. Another purpose of the current study was to determine how LMX and trustworthiness influence student reactions such as motivation to improve performance, accuracy ratings of the performance appraisal, and satisfaction of performance ratings. First, the hypothesized three factor model was compared to a one factor model in order to determine whether there were empirical grounds for treating the three proposed antecedents and moderators in the model tested here as distinct variables. In the three factor model, the factors include performance appraisal characteristics consisting of rater accountability, frequency of observation, justice perceptions and trust,

LMX consisting of affect, loyalty, professional respect and contribution, and trustworthiness consisting of benevolence, ability and integrity

Therefore the following hypotheses were tested:

Hypothesis 1: LMX and trustworthiness are positively related to performance appraisal characteristics

Hypothesis 2a: Performance appraisal system characteristics moderate the relationship between LMX and performance ratings such that high levels of system characteristics will strengthen the relationship between LMX and performance ratings.

Hypothesis 2b: Performance appraisal system characteristics moderate the relationship between trustworthiness and performance ratings such that high levels of system characteristics will strengthen the relationship between trustworthiness and performance ratings.

Hypothesis 3a: Performance ratings partially mediate the relationship between LMX and feedback reactions.

Hypothesis 3b: Performance ratings partially mediate the relationship between trustworthiness and feedback reactions.

Chapter 3

Method

Participants

The sample consisted of graduate students and graduate advisers from multiple departments of a large university. This sample was chosen because of the more controlled and stable environment of graduate school as compared to using supervisor-subordinate pairs in a more uncontrollable and unstable organizational environment. Using this sample was a good way for the researcher to test the hypothesized relationships without as many external factors possibly found in a work environment. For example, it is common for graduate students and advisers to have a strong, one-on-one relationship. Often graduate students attend a particular university with a specific adviser with mutual research interests in order for them to work together on targeted projects. In organizations, it is possible for employees to work with many supervisors on different projects. If this were the case, feedback from multiple supervisors would need to be collected; therefore testing these relationships would not be nearly as clean. Using this sample to test the model was a good first step in this type of research. Some of these relationships can be generalized to an organizational sample and research can test a similar model using an organizational sample in the future.

The sample size was 192 (96 advisees, 96 advisers at a 1:1 advisee to adviser ratio). The sample was gathered from various graduate programs throughout the university. All genders, races, and ethnicities were included in the sample.

Measures

Leader-Member Exchange

Both the student's and adviser's perception of the leader-member exchange relationship were assessed with the LMX-MDM scale (Liden & Maslyn, 1998) which consists of 12 items modified to fit the graduate student/adviser context. There are four dimensions of LMX that are measured in the LMX-MDM. An example item from the *contribution* scale (Cronbach's alpha ranging from .74 to .77) includes "I do not mind working my hardest for my adviser (student)." An example item from the *professional respect* scale (Cronbach's alpha ranging from .89 to .92) includes "I am impressed with my adviser's (student's) knowledge of his or her job." An example item from the *affect* scale (Cronbach's alpha ranging from .83 to .90) includes "I like my adviser (student) very much as a person." An example item from the *loyalty* scale (Cronbach's alpha ranging from .74 to .78) includes "My adviser defends my work actions to a superior even without complete knowledge of the issue in question" or "My student defends my decisions, even without complete knowledge of the issue in question." This measure will use a 5-point Likert scale ranging from *strongly disagree* to *strongly agree*.

Trustworthiness

Trustworthiness was looked at in terms of the student's perception of his or her adviser's ability, benevolence, and integrity (Schoorman et al., 1996). Perception of adviser's ability was assessed using a six item measure with a Cronbach's alpha ranging from .85 to .88. A sample item of ability is "My adviser is very capable of performing his or her job." Perception of adviser benevolence was assessed with a five item measure with a Cronbach's alpha ranging from .87 to .89. A sample item of benevolence is "My adviser is very concerned about my welfare." Integrity was assessed using a six item measure with a Cronbach's alpha ranging from .82 to .88. A sample item is "My adviser tried hard to be fair in dealings with others." This measure will use a 5-point Likert scale ranging from *strongly disagree* to *strongly agree*.

Trust in Performance Appraisal System

The trust in performance appraisal system measure was developed specifically for this study. Students were asked "Do you trust the performance appraisal system your department has in place?" Students indicated their responses on a 5-point Likert scale with anchors of *to a small extent* and *to a large extent*.

Frequency of Observation

Frequency of observation data was collected via direct self-reports from both graduate advisers (raters) and graduate students (ratees). In this context, it is helpful to consider observation as an "umbrella term that covers many methods that a supervisor might use to obtain information about subordinates" (Murphy & Cleveland, 1995). This

includes direct observations of the advisees' behavior, and observation of the results of that behavior (Murphy & Cleveland, 1995). According to Murphy and Cleveland (1995), much of what supervisors know about their subordinates' performance is the result of indirect observation. For example, indirect observation may include hearing rumors about the advisee, and reading descriptions of the advisees' behavior.

In the current study, participants were asked how frequently they were observed by their advisers and what types of observations were conducted (e.g. direct, indirect). Additionally, rater participants were asked questions such as 'How sufficient was the information you gathered on your student's performance in order to evaluate him or her?' The ratee participants were asked questions such as 'How sufficient was the information your adviser gathered on your performance in order to evaluate you?' Both rater and ratee measures used a 5-point Likert scale ranging from *excellent sufficiency* to *extremely poor sufficiency*.

Justice Perceptions

Because previous researchers have defined and measured distributive justice as agreement with and perceived fairness of the superior's performance rating (Folger, 1987; Greenberg, 1986), the current study used an altered version of the four item scale used in the Korsgaard and Roberson (1995) study. Items were rated on a 5-point Likert scale ranging from *strongly disagree* to *strongly agree*. Items included 1) the performance appraisal was fair, 2) I agree with my final rating, 3) I agree with the way my adviser rated my performance and 4) the performance appraisal fairly represented my past year's performance (Cronbach's alpha = .93).

Procedural justice was measured using seven items from Colquitt (2001). This scale was altered to reflect the performance appraisal context which is appropriate for the current study. Students indicated their responses on a 5-point Likert scale with anchors of *to a small extent* and *to a large extent*. The items assess the perceived fairness, appropriateness, and impartiality of the appraisal process. Sample items from this scale include “Have you been able to express your views and feelings during performance appraisal procedures” and “Have the performance appraisal procedures been applied consistently?” The internal consistency reliability is $\alpha = .78$.

Interactional justice was measured using nine items (4 interpersonal justice items, 5 informational justice items) from Colquitt (2001). This scale was altered to reflect the performance appraisal context which was appropriate for the current study. Students indicated their responses on a 5-point Likert scale with anchors of *to a small extent* and *to a large extent*. An example item from the interpersonal justice scale is “Has your adviser treated you with respect?” An example item from the informational justice scale is “Has your adviser explained the performance appraisal procedures thoroughly?” The internal consistency reliability of both interpersonal and interactional justice is $\alpha = .79$.

Rater Accountability

A measure of rater accountability was developed specifically for this study. The items included in this measure explore if advisers are currently held accountable for performance evaluations and what advisers believe to be the consequences of both downward and upward accountability. Advisers indicated their responses on a 5-point Likert scale ranging from *strongly disagree* to *strongly agree*. Sample items from the

downward accountability measure include “I feel accountable to my advisee for the completeness of my performance evaluation” and “The completeness of my performance evaluations is linked to my relationship with my advisee.” Sample items from the upward accountability measure include “I believe my department chair has read the evaluation I made” and “I believe my colleagues care about the evaluation I made.”

Performance

Performance was measured from both the adviser and student perspective. The performance measure was developed specifically for this study. Advisers were asked to rate performance based on 4 criteria: Student’s dependability, quality of work, participation in OCBs, and progress in terms of meeting milestones. Students rated themselves on the same criteria. Ratings were made on a 7-point Likert scale.

Appraisal Reactions

Participants completed a three item measure of satisfaction with the performance appraisal system specifically developed for this study. The measure uses a 5-point Likert response scale ranging from *strongly disagree* to *strongly agree* (‘Overall, I felt quite satisfied with the performance appraisal process my graduate program used’, ‘I felt quite satisfied with the form in which I received feedback from my adviser’, ‘I felt quite satisfied with the performance feedback my adviser gave to me’).

Motivation to improve performance was measured by a three item questionnaire adapted from Dorfman, Stephan, and Loveland (1986) referring to whether students will feel motivated to do their very best in the future, whether they are willing to put forth

effort beyond that normally expected, and the extent to which they will want to improve their future performance as a result of the performance appraisal. The items were rated on a 5-point Likert scale ranging from *strongly disagree* to *strongly agree* and combined to form a composite index of motivation (coefficient alpha = .69).

The extent to which employees perceived their appraisal as accurate was measured by a 7-item scale by Mayer and Davis (1995). Students indicated their responses on a 5-point Likert scale ranging from *strongly disagree* to *strongly agree*. A sample item is “The evaluation of what skills I have is pretty accurate.” Alphas for this scale range from .70 to .82.

Procedure

Because this study used adviser-student dyads, two surveys were developed: an adviser version and a graduate student version. The adviser version of the survey included the frequency of observation, rater accountability, LMX, and performance criteria measures. The student survey version included all measures except rater accountability. This survey was designed to be available to be taken at any point in time as long as the student had been evaluated in the past and had an adviser in mind as a referent for the LMX measure. Recruitment emails (see Appendix C and Appendix D) with the survey link were sent to 1500 advisers. If the adviser decided to participate, he or she would send the student recruitment letter with the survey link to one of his or her students.

Chapter 4

Results

There were a total of 96 graduate student-adviser pairs. The adviser sample consisted of 61 males and 35 females ranging from 24 to 69 years of age ($M = 48.87$, $SD = 10.33$). Eighty-one advisers were White, seven were White non-Hispanic, none were Hispanic, three were African-American, four were Asian-Pacific Islander and one participant marked 'other' for this item. Seventy-six percent of the advisers in the sample were tenured. The student sample consisted of 38 males and 57 females ranging from 23 to 62 years old ($M = 29.82$, $SD = 6.36$). Fifty-four students were White, sixteen were White, non-Hispanic, two were African-American, five were Hispanic, 13 were Asian-Pacific Islander and four marked 'other' for this item.

Descriptive Statistics

Means, standard deviations and reliabilities for the variables measured in this study are presented in Table 1. Overall, LMX was rated high by both students and advisers; however there appears to be a restriction of range issue. Neither the students nor the advisers used the full rating scale for this variable. A restricted range can result in a reduced correlation between adviser and student ratings of LMX. Student ratings of adviser trustworthiness were high; however the full rating scale of trustworthiness was not used resulting in a restriction of range as well. Ratings of trust in the overall performance appraisal system were high (above 3 on a 5-point scale) and the full range of the rating scale was utilized. Adviser and student ratings of frequency of observation

were also high and the full range of the rating scale was used. Although ratings of all four facets of performance and overall performance were generally high (above 3 on a 5-point scale), it could be argued that there was a restriction of range, particularly with the adviser ratings of performance. Advisers did not use the full rating scale for any of the performance facets. Additionally, students did not use the full rating scale for two of the four performance facets. A restricted range can result in a reduced correlation between adviser and student ratings of performance. The results of the descriptive statistics could, to some degree, be due to selection bias. For example, the adviser was asked to select a student he or she wanted to use as a referent for the survey. It is likely advisers chose students who they not only have a high LMX relationship with, but who are also high performers which may have led to this restriction of range in both of these measures.

Reliability of Scales

The reliabilities for each scale in the survey and the overall reliability of the measure were calculated. The overall Cronbach's alphas for the adviser survey and the student survey were .84 and .96 respectively. However, there were two scales that had a Cronbach's alpha below .70; student ratings of their own performance (.53) and student's measure of frequency of observation (.69).

There were less than 15 cases of missing data. Parallel analyses were run using mean substitution versus dropping missing data (pairwise) from the analysis and there were no major differences. Means, standard deviations, and reliabilities for all of the variables examined can be seen in Table 1 below.

Table 1

Means, Standard Deviations, and Reliabilities

	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>Reliability</i>
Adviser LMX	90	50.10	5.12	0.84
Student LMX	93	52.00	5.37	0.83
Adviser Frequency	94	16.98	2.39	0.81
Student Frequency	96	15.96	2.64	0.69
Accountability	94	13.59	2.89	0.65
Trustworthiness	93	74.59	7.71	0.90
Justice	87	83.23	11.99	0.94
Trust in Performance Appraisal System	96	3.13	0.78	(1 item measure)
Student Dependability	96	3.59	0.70	(1 item measure)
Student Quality	96	4.52	0.54	(1 item measure)
Student OCB	96	4.61	0.57	(1 item measure)
Student Timeliness	96	4.32	0.80	(1 item measure)
Student Overall Performance	96	17.05	1.71	0.53
Adviser Dependability	96	3.8	0.40	(1 item measure)
Adviser Quality	96	4.66	0.52	(1 item

				measure)
Adviser OCB	96	4.66	0.60	(1 item measure)
Adviser Timeliness	96	4.33	0.71	(1 item measure)
Adviser Overall Performance		17.45	1.69	0.73
Accuracy	95	14.21	2.84	0.72
Motivation	95	16.57	2.56	0.74
Satisfaction	96	16.88	2.62	0.79
Adviser Survey	96			0.84
Student Survey	96			0.96

Confirmatory Factor Analysis

Conducting confirmatory factor analysis (CFA) allowed the researcher to test the hypothesis that the set of observed variables loaded on the three factor structure hypothesized using knowledge of theory and empirical research. More specifically that trust in system, frequency of observation, rater accountability, and justice perceptions loaded onto the first factor (performance appraisal system characteristics), student and adviser ratings of LMX loaded on the second factor (LMX), and trustworthiness loaded onto a third factor. This 3 factor model could then be compared to a 1 factor model in

order to determine whether these three variables were sufficiently distinct to treat as different variables.

To test and compare the measurement models, two CFAs were conducted using LISREL to assess the fit of the current study's 3 factor model and to compare this fit to the fit of a simpler 1 factor model. The latent variables in the 3 factor model were LMX (student ratings, adviser ratings), student trustworthiness and performance appraisal system characteristics (rater accountability, frequency of observation, trust in system, justice perceptions). The comparative fit index (CFI) for the 3 factor model was .96 while the CFI for the 1 factor model was .87. The Root Mean Square Error of Approximation (RMSEA) for the 1 factor model was .13 and the RMSEA of the 3 factor model was .074. Models with an RMSEA of .10 or higher are considered to have poor fit (Bollen & Long, 1993); therefore the 1 factor model does not fit well. Taken together these results demonstrate that the 3 factor model fits well and that the variables included in this model are measuring more than one factor. Additionally, the CFA provides rationale for forming composite variables described in the next section. Although CFA was used to compare the 3 factor and the 1 factor model, the sample was not sufficient in size to obtain stable estimates of the structural parameters proposed in Figure 1 and Figure 2. These results suggest adviser and student results should be examined separately. On the other hand, aggregated results have the advantage of being more reliable.

Unit of Analysis

The purpose of the current study was to determine how rater accountability, frequency of observation, justice perceptions, and trust in the appraisal system impact the

relationship between LMX and performance ratings, and between trustworthiness and performance ratings. A second purpose of this study was to look at factors influencing feedback reactions (motivation to improve performance, accuracy, satisfaction) to performance appraisals. The survey allowed one-to-one student-adviser pairs to describe the performance appraisal system in terms of rater accountability, frequency of observation, justice perceptions and trust in the overall performance appraisal system.

Due to the fact that each student-adviser pair described the same system and the results of the above CFA demonstrated that the 3 factor model was a good fit, it was logical to combine their data to form a composite variable titled Performance Appraisal System Characteristics. Individual responses were transformed into z scores prior to forming a composite to give equal weight to student and adviser perceptions.

Two CFAs were conducted in order to compare the fit statistics of the 3 factor model with the variables LMX (student ratings, adviser ratings), student trustworthiness and performance appraisal system characteristics (rater accountability, frequency of observation, trust in system, justice perceptions) to a 1 factor model. Because the 3 factor model resulted in better fit, the three composite variables were formed for each of the three factors in the model. Due to the small sample size, it made sense to form these composites and run the analyses. Based on logic and past research examining these types of variables and demonstrating how they relate to each other, it was decided that the unit of analysis for each variable such as LMX, performance ratings and the performance appraisal system characteristics would be a combined score of both student and adviser data rather than running separate analyses for students and advisers.

Additionally, student and adviser performance ratings were combined for each facet of performance (quality, timeliness, OCBs, dependability) and for overall performance, again using combined z-scores. These facets of performance were one-item measures which do not have high reliability; therefore they were collapsed into the overall performance composite. Finally, the unit of analysis for LMX was also formed into a composite variable using combined z-scores of student and adviser LMX data. The concept of LMX is very much dependent on both the leader and the member and it should always be measured from both perspectives (Gerstner & Day, 1997). Research has indicated that LMX should be analyzed at the dyadic level and there appears to be basic agreement within the field regarding this position (Schriesheim, Castro, & Cogliser, 1999). A Pearson r correlation was used to describe the relationship between student and adviser ratings of leader-member exchange ($r(96) = .23, p < .05$). Pearson r correlations were conducted to look at the relationship between student self-ratings of performance and adviser ratings of performance. Data were collected on the four facets of performance rated by both students and their advisers: dependability ($r(96) = .09, n.s.$), timeliness ($r(96) = .31, p < .01$), organizational citizenship behaviors ($r(96) = .13, n.s.$) and quality of work ($r(96) = .12, n.s.$). Although some of the performance facets were not significantly correlated, student self-ratings of overall performance and adviser ratings of overall performance were significantly correlated ($r(96) = .25, p < .05$). All of the LMX and performance rating correlations are generally positively correlated, but they are not very high correlations. This is likely due to the restriction of range discussed at the beginning of this results section in both the LMX variables and the performance facets. The effect of a restricted range is a reduced correlation.

The correlations demonstrate that advisers and students do not agree very strongly on many of these variables (see Table 2 below). Additionally, the researcher ran two confirmatory factor analyses (CFAs) using the performance ratings data to determine the appropriate unit of analyses. A 2 factor model (student ratings of performance and adviser ratings of performance) was compared to a 1 factor model (an aggregate of the student and adviser ratings of performance). The CFI for the 2 factor model was .96 while the CFI for the 1 factor model was .85. The RMSEA for the 1 factor model was .12 and the RMSEA of the 2 factor model was .05. As stated previously, models with an RMSEA of .10 or higher are considered to have poor fit (Bollen & Long, 1993); therefore the 1 factor model does not fit well. Taken together these results demonstrate that the 2 factor model fits well and that the variables included in this model are measuring more than one factor.

The empirical case for looking at aggregated versus non-aggregated results is not clear; therefore the researcher replicated the analyses at three levels: adviser, student, and aggregated. The conclusions are the same at all three levels of analysis, and because the aggregated analyses are more reliable, the aggregated results will be described in the text and the adviser only and student only results will be presented in Appendices E through H.

Table 2

Correlations between Performance Appraisal System Characteristics, Leader-Member Exchange, Trustworthiness, Performance Ratings and Feedback Reactions (Individual Level)

Variables	1	2	3	4	5	6	7	8	9	10
1. Student Timeliness	X									
2. Student Quality	.26**	X								
3. Student Dependability	.22*	.23*	X							
4. Student OCBs	0.14	.28**	.26**	X						
5. Student Overall	.62**	.68**	.66**	.62**	X					
6. Adviser Timeliness	.31**	.23*	0.15	-0.12	.23*	X				
7. Adviser Quality	0.12	0.12	0.1	0.12	0.18	.43**	X			
8. Adviser Dependability	0.14	0.14	0.09	-0.02	0.13	.57*	.58*	X		
9. Adviser OCBs	0.19	0.17	0.12	0.13	.23*	.28**	.40**	.42**	X	
10. Adviser Overall	.25*	.25*	0.15	.25*	0.25*	0.75	.79**	.84**	.68**	X
11. Adviser LMX	0.10	0.12	.21*	.31**	.29**	0.2	.54**	.44**	.42**	.52**
12. Student LMX	0.05	0.04	-0.05	0.14	0.07	-0.02	0.13	0.16	.33**	0.2
13. Trustworthiness	0.03	0.02	-0.09	0.00	-0.02	-0.09	0.05	0.06	0.14	0.05
14. Trust in PA System	-0.03	0.19	0.13	0.06	0.14	0.06	0.00	0.01	-0.16	-0.03
15. Adviser Frequency	0.09	.27**	0.03	0.06	0.17	.23*	0.19	.25*	0.19	.28**
16. Student Frequency	.23*	.29**	0.08	0.07	.25*	-0.03	0.16	0.14	0.03	0.1
17. Accountability	-0.09	0.11	0.03	0.05	0.04	0.12	0.14	0.12	0.03	0.13
18. Justice	0.03	0.07	0.02	-0.06	0.02	0.04	0.02	0.13	0.08	0.09
19. Satisfaction	0.12	0.14	0.06	0.02	0.13	0.1	0.15	6	0.18	0.19
20. Motivation	.22*	.23*	0.05	0.13	.24*	0.02	0.04	0.1	0.13	0.1
21. Accuracy	0.22	0.17	0.17	0.07	0.19	0.08	-0.02	0.11	0.03	0.07

* $p < .05$; ** $p < .01$; *** $p < .001$

Variables	11	12	13	14	15	16	17	18	19	20	21
1. Student Timeliness											
2. Student Quality											
3. Student Dependability											
4. Student OCBs											
5. Student Overall											
6. Adviser Timeliness											
7. Adviser Quality											
8. Adviser Dependability											
9. Adviser OCBs											
10. Adviser Overall											
11. Adviser LMX	X										
12. Student LMX	.20*	X									
13. Trustworthiness	.21*	.79**	X								
14. Trust in PA System	0.10	.29*	.40***	X							
15. Adviser Frequency	.23*	0.14	0.15	0.17	X						
16. Student Frequency	0.14	.46***	.57***	.50*	.24**	X					
17. Accountability	.27**	0.11	0.1	0.18	.31**	0.06	X				
18. Justice	.21*	.55***	.58**	.46***	0.18	.56**	.20*	X			
19. Satisfaction	0.18	.59***	.63***	.57***	.28**	.64***	.27**	.76**	X		
20. Motivation	0.04	.58***	.47***	.43***	.21*	.52***	0.11	.47***	.60***	X	
21. Accuracy	0.02	0.14	0.14	.24*	0.08	.31**	0.13	.30**	.36**	.54***	X

* $p < .05$; ** $p < .01$; *** $p < .001$

Statistical Analyses

Hypothesis 1 states that LMX and trustworthiness are positively related to performance appraisal characteristics. Pearson r correlations were used to look at the relationship between LMX, trustworthiness, performance appraisal characteristics and overall performance. LMX and performance appraisal characteristics were positively correlated ($r(96) = .54, p < .001$). Trustworthiness and performance appraisal characteristics were also positively correlated ($r(96) = .61, p < .001$). It was also important to look at the relationship between these variables and the dependent variable of overall performance. However, LMX was the only variable that was significantly correlated to overall performance ($r(96) = .46, p < .001$). Table 3 below displays all correlations.

Table 3

Correlations between Performance Appraisal System Characteristics, Leader-Member Exchange, Trustworthiness, Performance Ratings and Feedback Reactions

Variables	1	2	3	4	5	6	7	8	9	10	11
1. Accountability	X										
2. Frequency	.25*	X									
3. Justice	.22*	.45**	X								
4. Trust in System	.18	.41**	.46**	X							
5. LMX	.24*	.40**	.49**	.54**	X						
6. Trustworthiness	.11	.46**	.58**	.61**	.54**	X					
7. Performance	.11	.33**	.07	.17	.46**	.02	X				
8. Satisfaction	.27**	.57**	.76**	.81**	.50**	.63**	.21*	X			
9. Motivation	.11	.45**	.47**	.52**	.40**	.48**	.20*	.62**	X		
10. Accuracy	.13	.24*	.30**	.32**	.10	.16	.15	.34**	.56**	X	
11. Performance Appraisal System Characteristics	.39**	.69**	.95**	.55**	.54**	.61**	.17	.81**	.52**	.32**	X

* $p < .05$; ** $p < .01$.

In order to test Hypothesis 2a, which states that performance appraisal system characteristics moderate the relationship between LMX and performance ratings such that high levels of system characteristics will strengthen the relationship between LMX and performance ratings, moderated multiple regression analyses were conducted. To test hypothesis 2b, which states that performance appraisal system characteristics moderate the relationship between trustworthiness and performance ratings such that high levels of

system characteristics will strengthen the relationship between trustworthiness and performance ratings, moderated multiple regression using were conducted. Each facet of performance was tested in addition to overall performance. Tables 4 through 8 below display the result of each moderated multiple regression analysis.

The first facet of performance tested was ratings of organizational citizenship behaviors. A significant R^2 emerged ($F(6, 89) = 7.95, p < .001, R^2 = .35, \text{Adjusted } R^2 = .31$; see Table 4). LMX was a significant predictor of ratings of performance for this facet ($\beta = .74, p < .01$). The performance appraisal system characteristic variable was not found to be a moderator. The R^2 value before the cross products were entered was .33 and after the cross products were entered R^2 was .35; therefore change in R^2 was .02.

Table 4

Moderated Multiple Regression: OCBs as Dependent Variable

<i>Variable</i>	<i>B</i>	<i>B</i>	<i>R²</i>	<i>ΔR²</i>
Step 1			0.30	
Trust	-0.05	-0.26**		
LMX	0.09	0.65*		
Step 2			0.33	0.03
PA Sys Char	-0.02	-0.23*		
Step 3			0.34	0.00
Sys*LMX	0.00	0.04		
Sys*Trust	0.00	-0.34		
Step 4			0.35	0.013
Sys*LMX*Trust	0.00	-0.46		

Note: $R = .59, R^2 = .35, \text{ and } \text{Adj. } R^2 = .31$

Note: $N = 95$

* $p < .05, **p < .001$

The result of the moderated multiple regression with timeliness as the facet of performance tested was non-significant ($F(6, 89) = .91$ *n.s.*; see Table 5) with an R^2 of .06 and a change in R^2 of .01.

Table 5

Moderated Multiple Regression: Timeliness as Dependent Variable

<i>Variable</i>	B	B	R^2	ΔR^2
Step 1			0.03	
Trust	-0.04	-0.16		
LMX	0.03	0.22		
Step 2			0.05	0.01
PA Sys Char	0.01	0.13		
Step 3			0.06	0.01
Sys*LMX	0.00	0.21		
Sys*Trust	0.00	-1.07		
Step 4			0.06	0.001
Sys*LMX*Trust	0.00	-0.27		

Note: $R = .24$, $R^2 = .06$, and $\text{Adj. } R^2 = -.01$

Note: $N = 95$

Quality of work was the second facet of performance tested. A significant R^2 also emerged ($F(6, 89) = 3.08$ $p < .01$; $R^2 = .17$, Adjusted $R^2 = .12$; see Table 6). LMX was a significant predictor of ratings of performance for this facet ($\beta = .45$, $p < .001$); however there was no moderation effect. The R^2 value before the cross products were entered was .17 and after the cross products were entered R^2 was .17; therefore change in R^2 was .00.

Table 6

Moderated Multiple Regression: Quality is the Dependent Variable

<i>Variable</i>	<i>B</i>	<i>B</i>	<i>R</i> ²	ΔR^2
Step 1			0.16	
Trust	-0.04	-0.21		
LMX	0.06	0.47*		
Step 2			0.17	0.007
PA Sys Char	0.01	0.11		
Step 3			0.17	0.00
Sys*LMX	0.00	0.07		
Sys*Trust	0.00	-0.34		
Step 4			0.17	0.004
Sys*LMX*Trust	0.00	-0.25		

Note: $R = .42$, $R^2 = .17$, and $\text{Adj. } R^2 = .12$

Note: $N = 95$

* $p < .001$

Next the facet of dependability was tested. A significant R^2 emerged ($F(6, 89) = 4.49$, $p < .05$; $R^2 = .23$, $\text{Adjusted } R^2 = .18$; see Table 7). LMX and trustworthiness were both significant predictors of dependability ($\beta = .42$, $p < .05$ and $\beta = -.54$, $p < .05$). The performance appraisal system characteristic variable was not found to be a moderator. The R^2 value before the cross products were entered was .17 and after the cross products were entered R^2 was .23; therefore change in R^2 was .06. This change in R^2 is substantial and with more powerful techniques, it is possible that it would reach significance.

Table 7

Moderated Multiple Regression: Dependability as the Dependent Variable

<i>Variable</i>	<i>B</i>	<i>B</i>	<i>R</i> ²	ΔR^2
Step 1			0.17	
Trust	-0.05	-0.27*		
LMX	0.06	0.47**		
Step 2			0.17	0.01
PA Sys Char	0.01	0.116		
Step 3			0.21	0.05
Sys*LMX	0.00	0.11		
Sys*Trust	0.00	-1.72		
Step 4			0.23	0.02
Sys*LMX*Trust	0.00	0.54		

Note: $R = .48$, $R^2 = .23$, and $\text{Adj. } R^2 = .18$

Note: $N = 95$

* $p < .05$, ** $p < .001$

Lastly, overall performance ratings were tested. A significant R^2 emerged ($F(6, 89) = 5.76$ $p < .001$; $R^2 = .28$; Adjusted $R^2 = .23$; see Table 8). LMX and trustworthiness were both significant predictors of overall performance ($\beta = .60$, $p < .001$ and $\beta = -.41$, $p < .05$). Performance appraisal system characteristics did not moderate any relationships. The R^2 value before the cross products were entered was .26 and after the cross products were entered R^2 was .28; therefore change in R^2 was .02.

Table 8

Moderated Multiple Regression: Overall Performance as the Dependent Variable

<i>Variable</i>	<i>B</i>	<i>B</i>	<i>R</i> ²	ΔR^2
Step 1			0.26	
Trust	-0.18	-.30*		
LMX	0.24	0.61**		
Step 2			0.26	0.001
PA Sys Char	0.01	0.04		
Step 3			0.28	0.02
Sys*LMX	0.00	0.15		
Sys*Trust	0.00	-1.18		
Step 4			0.28	0.001
Sys*LMX*Trust	0.00	-0.28		

Note: $R = .53$, $R^2 = .28$, and $\text{Adj. } R^2 = .23$

Note: $N = 95$

* $p < .01$, ** $p < .001$

Overall, change in R^2 ranged from 0 to .06 depending on the performance facet. More specifically, performance appraisal system characteristics explained 6% more of the variance in dependability than just LMX and trustworthiness alone. Performance appraisal system characteristics explained 2% more of the variance in OCBs than just LMX and trustworthiness alone. Performance appraisal system characteristics explained 2% more of the variance in overall performance than just LMX and trustworthiness alone. These results give some supporting evidence that performance system appraisal characteristics such as justice perception, frequency of observation, rater accountability, and trust in the performance appraisal system can have an impact on performance ratings.

Methods of Mediation

The following section describes alternative methods of mediation. Although the Baron and Kenny (1986) method of testing for mediation is classic, it is somewhat controversial because it relies on significance testing. However, there are now over a dozen methods used to test for mediation; however Mackinnon, Lockwood, Hoffman, West, and Sheets (2002) discuss three broad approaches to mediation testing. They combine the causal step method of Baron and Kenny with the differences in coefficients approach (DCA), and product of coefficients approach (PCA) or the Sobel formula. The causal step method establishes conditions, but does not provide evidence for causal effects. Additionally, this method relies solely on significance testing because it has no estimate of effect size. DCA provides an estimate of the mediating variable's effect size and its standard error. Lastly, PCA uses the Sobel formula in order to test for significance. More specifically, the Sobel formula divides the mediating variable's effect size by its standard error and compares this value to a normal distribution. In Mackinnon et al.'s (2002) study, the researchers conducted 14 Monte Carlo simulations with 14 mediation models and the PCA and DCA methods had the most accurate Type 1 error rates and greatest statistical power.

Despite its limitation, the current study used the Baron and Kenny (1986) approach to mediation because it is the most commonly used and most frequently cited test of mediation throughout the psychology literature. Additionally, the alternative methods have limitations as well. For example, the Sobel test assumes that the sampling distribution is normal. The current study had a small sample size making it difficult to

reach this assumption. Sample size also prevented the research from conducting the mediation analysis using SEM.

Mediated multiple regression was used to test hypothesis 3a and 3b which state that performance ratings partially mediate the relationship between LMX and feedback reactions (motivation to improve performance, satisfaction, accuracy) and performance ratings partially mediate the relationship between trustworthiness and feedback reactions. Baron and Kenny (1986) proposed a four step approach in which several regression analyses are conducted and significance of the coefficients is examined at each step. Step 1 consists of conducting a simple regression analysis with X predicting Y. In Step 2 a simple regression analysis is conducted with X predicting M. Step 3 entails conducting a simple regression analysis with M predicting Y. Lastly, Step 4 consists of conducting a simple regression analysis with X and M predicting Y. If one or more of the relationships in Steps 1 through 3 are non-significant, researchers usually conclude that mediation is not possible or likely (Baron & Kenny, 1986). Assuming there are significant relationships from Steps 1 through 3, one proceeds to Step 4. In the Step 4 model, some form of mediation is supported if the effect of M remains significant after controlling for X. If X is no longer significant when M is controlled, the finding supports full mediation (Baron & Kenny, 1986). If X is still significant, the finding supports partial mediation (Baron & Kenny, 1986) which was hypothesized in the current study.

The first mediation tested was overall performance ratings mediating the relationship between LMX and the student feedback reaction satisfaction. The relationship between LMX and satisfaction, LMX and overall performance ratings, and performance ratings and satisfaction are $\beta = .50, p < .001$, $\beta = .55, p < .001$ and $\beta = .21, p$

< .05 respectively. Based on these significant results, Step 4 was completed. The simple regression for Step 4 resulted in a continued significant relationship between LMX and satisfaction ($\beta = .02, p < .001$); but the mediator, overall performance, was not significantly related to satisfaction ($\beta = .51, n.s.$). The relationship between LMX and satisfaction had an R^2 of .248. When the mediator was included, the R^2 value was .249 making change in R^2 .001. See Table 9 for R^2 values.

The second mediation tested was overall performance ratings mediating the relationship between LMX and the student feedback reaction motivation. The relationship between LMX and motivation, LMX and overall performance ratings and performance ratings and motivation are $\beta = .40, p < .001$, $\beta = .45, p < .001$ and $\beta = .20, p < .05$ respectively. Based on these significant results, Step 4 was completed. The simple regression for Step 4 resulted in a continued significant relationship between LMX and motivation ($\beta = .39, p < .001$); but the mediator, overall performance, was not significantly related to satisfaction ($\beta = .03, n.s.$). The relationship between LMX and motivation had an R^2 of .16. When the mediator was included, the R^2 value was .161 making change in R^2 .001. This could indicate that overall performance ratings partially mediate the relationship between LMX and motivation. Table 10 displays R^2 values.

The third mediation tested was overall performance ratings mediating the relationship between LMX and the student feedback reaction accuracy. The relationship between LMX and accuracy, LMX and overall performance ratings and performance ratings and accuracy are $\beta = .10, n.s.$, $\beta = .45, p < .001$ and $\beta = .15, n.s.$ respectively. Based on these non-significant results, Step 4 was not completed. This indicates that

overall performance ratings do not mediate the relationship between LMX and accuracy. The relationship between LMX and accuracy had an R^2 of .01.

The fourth mediation tested was overall performance ratings mediating the relationship between trustworthiness and the student feedback reaction satisfaction. The relationship between trustworthiness and satisfaction, trustworthiness and overall performance ratings and performance ratings and satisfaction are $\beta = .63, p < .001$, $\beta = .02, n.s.$ and $\beta = .21, p < .05$ respectively. Based on these non-significant results, Step 4 was not completed. This indicates that overall performance ratings do not mediate the relationship between trustworthiness and satisfaction. The relationship between trustworthiness and satisfaction had an R^2 of .40. See Table 11 for R^2 values.

The fifth mediation tested was overall performance ratings mediating the relationship between trustworthiness and the student feedback reaction motivation. The relationship between trustworthiness and motivation, trustworthiness and overall performance ratings and performance ratings and motivation are $\beta = .48, p < .001$, $\beta = .02, n.s.$ and $\beta = .20, p < .05$ respectively. Based on these non-significant results, Step 4 was not completed. This indicates that overall performance ratings do not mediate the relationship between trustworthiness and motivation. The relationship between trustworthiness and satisfaction had an R^2 of .23. Table 12 displays R^2 values.

The last mediation tested was overall performance ratings mediating the relationship between trustworthiness and the student feedback reaction accuracy. The relationship between trustworthiness and accuracy, trustworthiness and overall performance ratings and performance ratings and accuracy are $\beta = .16, n.s.$, $\beta = .02, n.s.$ and $\beta = .15, n.s.$ respectively. Based on these non-significant results, Step 4 was not

completed. This indicates that overall performance ratings do not mediate the relationship between trustworthiness and accuracy. The relationship between trustworthiness and satisfaction had an R^2 of .02.

Due to the low power of the Baron and Kenny (1986) technique possibly explaining the failure to find mediation, the researcher compared the R^2 values of the mediator (performance ratings) regressed onto Y (feedback reactions), X (LMX or trustworthiness) regressed onto Y, and X and M regressed onto Y. This method applied to four out of the six mediations. This method could not be used to test if performance ratings mediate the relationship between LMX and accuracy because neither LMX nor performance ratings significantly related to accuracy. The researcher also could not use this approach when testing if performance ratings mediate the relationship between trust and accuracy because neither trustworthiness nor performance ratings were significantly related to accuracy. Tables 9 through 12 present the relevant R^2 values.

Table 9

Performance Ratings Mediating the Relationship between LMX and Satisfaction

<i>Variable</i>	R^2
LMX	0.25
Performance Ratings	0.04
LMX + Performance Ratings	0.25

Table 10

Performance Ratings Mediating the Relationship between LMX and Motivation

<i>Variable</i>	R^2
LMX	0.16
Performance Ratings	0.04
LMX + Performance Ratings	0.16

Table 11

Performance Ratings Mediating the Relationship between Trustworthiness and Satisfaction

<i>Variable</i>	R^2
Trustworthiness	0.40
Performance Ratings	0.04
Trustworthiness + Performance Ratings	0.44

Table 12

Performance Ratings Mediating the Relationship between Trustworthiness and Motivation

<i>Variable</i>	R^2
Trustworthiness	0.23
Performance Ratings	0.04
Trustworthiness + Performance Ratings	0.26

Based on these comparisons, it appears that performance ratings explain additional variance in the relationship between trustworthiness and satisfaction and trustworthiness and motivation. The results displayed in Tables 9 through 12 make it evident that no strong mediator effect would be found with more powerful techniques. These results further suggest that there is no advantage to considering alternative

mediation approaches to the Baron and Kenny (1986) method. Although there is some evidence of partial mediation, this may not mean more than the fact that X (LMX or trustworthiness), Y (feedback reactions), and M (performance ratings) are all positively correlated; therefore this aspect of the model is not supported.

Chapter 5

Discussion

Performance appraisals are conducted to evaluate what individuals do on the job (Brannick & Levine, 2002) and can have a number of purposes such as aiding with administrative decisions and providing feedback and development (Murphy & Cleveland, 1995). It is important to study performance appraisals in depth because this process is central to many personnel decisions (Landy, Barnes, & Murphy, 1978) and can greatly shape an organization. Much research on this topic has been conducted in the past because of the pivotal role of performance appraisal in human resource management (Mayer & Davis, 1999). However, most performance appraisal research has focused on aspects of performance appraisal such as rater errors and accuracy of ratings.

It has become clear to researchers and practitioners alike that performance appraisal takes place in a social context which plays a major role in the effectiveness of such appraisals and how participants react to the performance appraisal process (Bretz, Milkovich, & Read, 1992). However, there has not been sufficient emphasis on employee reactions to performance appraisal, which has been referred to as “neglected criteria” by some researchers (Murphy & Cleveland, 1995, p.310), regardless of the fact that reports suggest that many employees are considerably dissatisfied with the appraisal process in their organization (Elicker, Levy, & Hall, 2006). Importantly, Jacobs, Kafry, and Zedeck (1980) identified performance appraisal reactions such as satisfaction, accuracy and motivation to improve performance as a criterion of performance appraisal effectiveness.

Unfortunately, a survey of Fortune 100 firms revealed that only one-third of the organizations surveyed actually conduct attitude surveys to assess these types of variables (Thomas & Bretz, 1994).

This study expands on performance appraisal research and contributes to the field by including social context variables into the performance appraisal system model. First, the study examined how performance appraisal system characteristics (rater accountability, frequency of observation, justice perceptions, and trust in the appraisal system) impact the relationship between LMX and performance ratings, and between trustworthiness and performance ratings. This is important because past research has not empirically tested these relationships to determine if they do indeed exist.

Secondly, the study examined factors influencing feedback reactions (motivation to improve performance, accuracy, satisfaction) to performance appraisals. The study emphasizes not only the importance of feedback reactions in regards to performance, but what particular factors shape these reactions such as performance ratings, trustworthiness and LMX. Due to individual differences, there must be alternate factors related to positive and negative reactions to performance appraisals other than the feedback itself. This is important because past research has demonstrated that individual performance and overall organizational performance are impacted by factors such as employee reactions. For example, justice perceptions have been found to be related to organizational citizenship behaviors, employee attitudes, and retaliation behaviors (Erdogan et al., 2001). Although evidence of a strong and direct relationship between employee satisfaction and individual performance is not well established (Iaffaldando & Muchinsky, 1985), this relationship has been documented at the organizational level of

analysis. Ostroff (1992) looked at the relationship between employee satisfaction and organizational performance with a sample of teachers. Both correlation and regression analysis supported this relationship; as employee satisfaction increased so did organizational performance (Ostroff, 1992). Interestingly, to date there has not been research conducted around social context variables and the performance appraisal process, although it has been a topic of theoretical papers (Erdogan, 2002) and future research discussions.

The model tested many important relationships relevant to the performance appraisal context such as the relationship between LMX, trustworthiness and performance appraisal characteristics. Supporting the first hypothesis, LMX and performance appraisal characteristics were strongly related in a positive direction; as LMX scores increased, so did ratings of performance appraisal characteristics. Trustworthiness and performance appraisal characteristics were also strongly related in a positive direction; as ratings of trustworthiness increased so did ratings of performance appraisal characteristics. These two findings highlight the importance of all of these variables because organizations should want high ratings of these performance appraisal characteristics in addition to high ratings of LMX and trustworthiness.

It was also important to look at the relationship between LMX, trustworthiness, and performance appraisal characteristics and the dependent variable overall performance. The results revealed that LMX was the only variable of the three that was strongly and positively related to overall performance. Again, the non-significant findings may be due to the small sample size used in this study. However, as ratings of LMX increased, so did overall performance. This last relationship demonstrates and

confirms the findings of past LMX research that LMX is important to overall performance. Interestingly, the relationship between student and adviser ratings of LMX was found to be small. This suggests that students and advisers have different ideas of how they rate each other on elements such as professional respect, contribution and affect (the three facets of LMX). These findings are consistent with previous findings in LMX research such as a meta-analysis conducted by Gerstner and Day (1997) which resulted in a .29 correlation between leader and member ratings of LMX. Of course, in their conclusion, these authors suggest that leader-member agreement be assessed using longitudinal designs (Gerstner & Day, 1997).

Related to this, the relationship between student and adviser ratings of overall performance was found to be rather small, indicating that students have a different idea of how they are performing than their advisers do. More specifically, the relationship between student and adviser ratings of timeliness was the only facet of performance that resulted in a moderate effect size. Dependability, OCBs and quality of work did not have impressive results. It is important to find the root cause of these less than robust relationships. Goals and performance standards should be mutually understood in both a school and a workplace setting. It is possible that some graduate programs at this university do not clearly state what is expected of graduate students. Specifically, it appears that students and advisers are of mutual understanding when it comes to timelines, but not so with other important aspects of performance. Graduate programs likely communicate a timeline to students at the onset of the program; however the expectations of the other facets of performance may vary from adviser to adviser making things less clear for students. This could explain why the relationship between student

self-ratings of performance and adviser ratings of performance did not correlate strongly. It would be interesting to see how these results would differ in the workplace. Successful onboarding programs in conjunction with initial and frequent discussions with supervisors about goals, performance standards, and expectations should lead to a more mutual understanding of performance.

Through the moderation analyses, it was discovered that LMX is a significant predictor of organizational citizenship behaviors, quality of work, and overall performance. This means that the higher the LMX rating, the more likely the student's rating of overall performance, OCBs, and quality of work will be high. Trustworthiness was also found to be a significant predictor of dependability and overall performance. This means that the higher the trustworthiness rating, the more likely the student's rating of overall performance and dependability will be high. Although there were no significant findings in terms of performance appraisal system characteristics moderating the relationship between LMX and trustworthiness and the facets of performance and overall performance, this may be due to the small sample size used in the study. Looking at change in R^2 , which ranged from 0 to .06 depending on the performance facet, one might conclude that with a larger sample, significant findings may have been discovered. More specifically, performance appraisal system characteristics explained 6% more of the variance in dependability than just LMX and trustworthiness alone. To support this point, a power analysis was conducted and revealed that in order to detect a change in R^2 of .06; the researcher should have power of about .68. This change in R^2 (.06) can be considered a small to medium effect size which would have been detected in a study with a slightly larger N; therefore there is some evidence of a small moderator. These results

give some supporting evidence that performance system appraisal characteristics such as justice perception, frequency of observation, rater accountability, and trust in the performance appraisal system can have an impact on performance ratings.

Using the Baron and Kenny (1986) approach to mediation, it was found that overall performance ratings partially mediated the relationship between LMX and two of the three student feedback reactions to performance appraisal --satisfaction and motivation. This means that ratings of overall performance account for a portion of the relationship between LMX and satisfaction and LMX and motivation. Said differently, the relationship between LMX and student feedback reactions of performance appraisals is apparently a partially indirect one. These results suggest that LMX ratings become integrated with ratings of overall performance, which in turn influence students' ultimate feedback reactions to performance appraisals. It is important to mention here that partial mediation really means that the independent, dependent and mediating variables are all positively correlated; therefore this component of the model is not supported. Additionally, due to the fact that Tables 9 through 12 make it apparent that there is no strong mediator effect, there is no benefit to contemplating alternatives to the Baron and Kenny (1986) approach to mediation.

Limitations and Future Research

The first limitation in the current study was the size of the sample (N = 96). Although it was large enough to conduct analyses such as confirmatory factor analysis, moderated multiple regression and mediated multiple regression, it was not large enough to give the researcher enough power or the probability that the test will reject the null

hypothesis when the null hypothesis is false. The sample size was also not large enough to use SEM to test the hypotheses. Although the researcher contacted over 1500 graduate advisers in a variety of graduate programs across the university via the adviser recruitment email (see Appendix C), the response rate was only about 6.5 percent. Additionally, the sample used may not be generalizable to a workplace setting. Although this study attempted to capture the unique one-to-one relationship between a graduate student and his or her adviser, it is possible that a supervisor-subordinate relationship is not identical. While some organizations may have a more traditional, one-on-one supervisor-subordinate structure, many organizations have evolved into a structure where employees work with multiple supervisors and a 360 degree feedback approach is more effective for performance appraisal. However, testing the three factor model with a graduate student-adviser population and finding that it was a good fit provides a foundation that this model should be tested in the workplace. Future studies can test this model in an organization where performance appraisals are conducted. The challenge to conducting a study like this is not only finding an organization willing to participate in what it may consider to be sensitive research, but finding an organization that has one-to-one supervisor-subordinate relationships. If this is not the case and employees are supervised by multiple individuals such as in a matrixed organization, the model would need to be modified. It would be interesting to look at these variables in a matrixed organization and see if they have similar effects on performance and performance appraisal feedback reactions.

In addition to the limited sample size and the generalizability of the sample, some of the measures used in this research did not have adequate reliabilities. For

example, student self-ratings of performance and the frequency of observation measure both had Cronbach's alpha values of less than .70. In addition, measures of performance and trust in system were developed by the researcher and included very few items. Moreover, some variables in the tested model have both complex and numerous definitions throughout the literature. For example, trust has been defined as a multidimensional construct with components such as integrity, benevolence, and ability. It is unclear if the scores on these three dimensions of trust should form a composite of overall trust, or if a sole trust measure should have been used.

Another limitation to consider is the level of analysis. Most research on this topic only considers individual performance ratings, and individual reactions to these ratings. With the increase of organizations utilizing work teams, the question of team performance appraisal leaves room for future research using these same variables measured at the team level. Again, this type of research would call for a modification of the model used in the current study.

Lastly, some may argue that because the data were collected from only one source for some variables and all at the same time through one survey, there is potential for common methods bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Although inflation due to common method variance might be a concern, this problem is less of a threat than is commonly believed (Spector, 2006)

Practical Implications and Conclusions

This study has practical implications for both organizations and applied research. As the literature demonstrates, performance appraisal is a frequently occurring process

central to many personnel and human resource decisions in organizations. Unfortunately, performance appraisal systems themselves are flawed. Although industrial and organizational psychologists have researched psychometric issues for decades, they have often ignored the context in which performance appraisal takes place. This study suggests the importance of rater accountability, frequency of observation, LMX and trust on employee reactions to performance appraisal systems, which have often been found to be negative. The discovery of the impact of variables such as these which may change negative reactions to performance appraisal systems into positive reactions may help an organization's overall effectiveness and success. It would be beneficial for organizations to concentrate on addressing these types of performance appraisal system characteristics and possibly revamp their performance appraisal process. Small changes in the performance appraisal system may lead to large changes in the way employees think about this process and perform their job.

Organizations can easily make adjustments to performance appraisal system characteristics. For example, supervisors can be held more accountable for their ratings. They can have one-on-one meetings with their subordinates to discuss the performance appraisal ratings in detail by providing specific behavioral examples with dates associated with them. Having to provide these behavioral examples will naturally increase frequency of observation. Supervisors will observe their employees as often as they can in anticipation of the one-on-one performance appraisal meetings in the future. Supervisors can also be held upwardly accountable. The supervisor's superior can review the performance appraisals and ensure that each employee is being rated with care. They can look for common rater errors such as halo ratings and central tendency. There can be

consequences for supervisors who do not take the time to properly appraise their employees. This study demonstrated that performance appraisal system characteristics are related. Although this does not mean that one variable causes the other, it is true that there is an association between these variables. It is possible that increasing frequency of observation and upward and downward accountability will increase employee perceptions of justice and trust in the overall performance appraisal system characteristics. Future research can use these ideas to test the effects of changes in performance system characteristics. An organization with a mediocre performance appraisal system, or one that results in negative reactions from employees, can attempt to revamp the system and measure employee reactions.

This study took years of theoretical speculation about the social context variables of performance appraisal systems and tested it empirically using graduate students and advisers. Specifically, we now see the importance of LMX on satisfaction and motivation to improve performance. Moreover, although there were not significant findings in the moderation analyses likely due to the small sample size, small effect sizes were discovered. This indicates that performance appraisal system characteristics such as justice perceptions, trust in system, rater accountability and frequency of observation could have an influence on particular facets of performance and overall performance. This study is a good first step in testing these important and impactful relationships and should be researched further using larger samples and samples from an organization.

BIBLIOGRAPHY

- Argyris, C. (1964). *Integrating the Individual and the Organization*. New York: Wiley.
- Avolio, B.J. & Bass, B.M. (1988). Transformational leadership, charisma and beyond. In J.G Hunt, B.R. Balaga, H.P Bachler, & C. Schrisheim (Eds.), *Emerging Leadership Vista* (pp. 29-50). Pergamon Press: Emosford, NY.
- Avolio, B.J., Bass, B.M., & Jung, D.I. (1999). Reexamining the components of transformational and transactional leadership using the Multifactor Leadership Questionnaire. *Journal of Organizational and Occupational Psychology*, 72, 441-462.
- Baron, R.M., & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Bollen, K.A. & Long, S.J. (Eds). (1993). *Testing Structural Equation Models*. Newbury Park: Sage Publications.
- Borman, W.C. & Motowidlo, S.J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt and W.C. Borman (Eds.), *Personality Selection* (pp. 71-98). San Francisco: Jossey-Bass.
- Bretz, R.D., Milkovich, G.T., & Read, W. (1992). The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of Management*, 18, 321-352.
- Brannick, M.T & Levine, E. (2002) *Job Analysis: Methods, Research, and Applications for Human Resource Management in the New Millennium*. Erlbaum Publications.

- Cawley, B.D., Keeping, L.M., & Levy, P.E. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of field investigations. *Journal of Applied Psychology, 83*, 615-633.
- Colquitt, J.A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*, 386-400.
- Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*, 425-445.
- Colquitt, J.A., Scott, B.A., & Lepine, J.A. (2007). Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology, 92*, 909-927.
- Curtis, A.B.; Harvey, R.D.; & Ravden, D. (2005). Sources of political distortions in performance appraisals: Appraisal purpose and rater accountability. *Group Organizational Management, 30*, 42-60.
- Dansereau, F., Jr., Graen, G., & Haga, W.J. (1975). A vertical dyad linkage approach to leadership within formal organizations: A longitudinal investigation of the role making process. *Organizational Behavior and Human Performance, 13*, 46-78.
- Dienesch, R. M., & Liden, R. C. (1986). Leader-member exchange model of leadership: A critique and further development. *Academy of Management Review, 11*, 618-634.
- Dirks, K.T. & Ferrin, D.L. (2001). The role of trust in organizational settings. *Organizational Science, 12*, 450-467.

- Dirks, K.T. & Ferrin, D.L. (2002). Trust in leadership: Meta-analytic findings and implications for research and practice. *Journal of Applied Psychology, 87*, 611-628.
- Dorfman, P. W., Stephan, W. G., & Loveland, J. (1986). Performance appraisal behaviors: Supervisor perceptions and subordinate reactions. *Personnel Psychology, 39*, 579-597.
- Elicker, J.D., Levy, P.E. & Hall, R.J. (2006). The role of leader-member exchange in the performance appraisal process. *Journal of Management, 32*, 531-551.
- Engle, E.M. & Lord, R.G. (1997). Implicit theories, self-schema, and leader-member exchange. *Academy of Management Journal, 40*, 988-1010.
- Erdogan, B., Kraimer, M.L., & Liden R.C. (2001). Procedural justice as a two dimensional construct: An examination in the performance appraisal context. *Journal of Applied Behavioral Science, 37*, 205-222.
- Erdogan, B. (2002). Antecedents and consequences of justice perceptions in performance appraisals. *Human Resource Management Review, 12*, 555-578.
- Farr, J.L. & Jacobs, R. (2006). Trust us: New Perspectives on performance appraisal.
- Folger, R. & Konovsky, M.A. (1989). Effects of procedural and distributive justice on reactions to pay raise decisions. *Academy of Management Journal, 32*, 115-130.
- Folger, R., Konovsky, M.A., & Cropanzano, R. (1992). A due process metaphor for performance appraisal. *Research in Organizational Behavior, 14*, 129-177.
- Gardner, W.L. & Avolio, B.J. (1998). The charismatic relationship: A dramaturgical perspective. *Academy of Management Review, 23*, 32-58.

- Gerstner, C.R. & Day, D.V. (1997). Meta-analytic review of leader-member exchange theory: Correlates and construct issues. *Journal of Applied Psychology*, 82, 827-844.
- Graen, G. B., & Schiemann, W. (1978). Leader-member agreement: A vertical dyad linkage approach. *Journal of Applied Psychology*, 63, 206-212.
- Graen, G.B. & Uhl-Bien, M. (1995). Relationship-based approach to leadership: Development of leader member exchange (LMX) theory of leadership over 25 years. *Leadership Quarterly*, 6, 219-247.
- Greenberg, J. (1986). Organizational performance appraisal procedures: What makes them fair? In R.J. Lewicki, B.H. Sheppard, & B.H. Bazerman (Eds.), *Research on negotiation in organizations*, 1: 25-41. Greenwich, CT: JAI.
- Harris, M. M., Smith, D. E., & Champagne, D. (1995). A field study of performance appraisal purpose: Research- versus administrative-based ratings. *Personnel Psychology*, 48, 151-160.
- Iaffaldano, M.T., & Muchinsky, P.M. (1985) Job satisfaction and job performance: A meta-analysis. *Psychological Bulletin*, 97, 251-273.
- Jacobs, R., Kafry, D., & Zedeck, S. (1980). Expectations of behaviorally anchored rating scales. *Personnel Psychology*, 33, 595-640.
- Jung, D.I. & Avolio, B.J. (2000). Opening the black box: An experimental investigation of the mediating effects of trust and value congruence on transformational and transactional leadership. *Journal of Organizational Behavior*, 21, 949-964.
- Keeping, L.M. & Levy, P.E. (2000). Performance appraisal reactions: Measurement, modeling, and method bias. *Journal of Applied Psychology*, 85, 708-723.

- Keeping, L. M., Makiney, J. D., Levy, P. E., Moon, M., & Gillette, L. M. (1999). Self ratings and reactions to feedback: It's not how you finish, but where you start. In R. A. Noe (Chair), *New approaches to understanding employees' affective and behavioral responses to multi-rater feedback systems*. Symposium conducted at the 59th annual meetings of the Academy of Management, Chicago.
- Klimoski, R., & Inks, L. (1990). Accountability forces in performance appraisal. *Organizational Behavior and Human Decision Process*, 45, 194-208.
- Konovsky, M. A., & Cropanzano, R. (1991). Perceived fairness of employee drug testing as a predictor of employee attitudes and job performance. *Journal of Applied Psychology*, 76, 698-707.
- Korsgaard, M.A. & Roberson, L. (1995). Procedural justice in performance evaluation: The role of instrumental and non-instrumental voice in performance appraisal discussions. *Journal of Management*, 21, 657-669.
- Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin*, 87, 72-107.
- Landy, F.J., Barnes, J.L., & Murphy, K.R. (1978). Correlates of perceived fairness and accuracy of performance evaluation. *Journal of Applied Psychology*, 63, 751-754.
- Lawler, E.E. (1967). The multi-trait multi-rater approach to measuring managerial job performance. *Journal of Applied Psychology*, 51, 369-381.
- Lerner, J.S. & Tetlock, P.E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125, 255-275.
- Levy, P.E. & Williams, J.R. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management*, 30, 881-905.

- Liden, R. C., & Maslyn, J. M. (1998). Multidimensionality of leader-member exchange: An empirical assessment through scale development. *Journal of Management, 24*, 43-72.
- Liden, R.C., Erdogan, B., Wayne, S.J., & Sparrowe, R.T. (2006). Leader-member exchange, differentiation, and task interdependence: Implications for individual and group performance. *Journal of Organizational Behavior, 27*, 723-746.
- Mackinnon, D.P., Lockwood, C.M., Hoffman, J.M., West, S.G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83-104.
- Mayer, R.C. & Davis, J.H. (1999). The effect of the performance appraisal system on trust for management: A field quasi-experiment. *Journal of Applied Psychology, 84*, 123-136.
- Mayer, R.C., Davis, J.H., & Schoorman, F.D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*, 709-734.
- McFarlin, D.B. & Sweeney, P.D. (1992). Distributive and procedural justice as predictors of satisfaction and personal and organizational outcomes. *Academy of Management Journal, 35*, 626-637.
- Mero, N.P. & Motowidlo, S.J. (1995). Effects of rater accountability on the accuracy and the favorability of performance ratings. *Journal of Applied Psychology, 80*, 517-524.
- Moorman, R. H. (1991). Relationship between organizational justice and organizational citizenship behaviors: do fairness perceptions influence employee citizenship? *Journal of Applied Psychology, 76*, 845–855.

- Murphy, K.R. & Cleveland, J.N. (1995). *Understanding Performance Appraisal: Social, Organizational, and Goal-based Perspectives*. Sage: Thousand Oaks, CA.
- Organ, D.W. (1988). Organizational citizenship behavior: The good soldier syndrome. Lexington, MA: Lexington.
- Organ, D.W. & Ryan, K. (1995). A meta-analytical review of attitudinal and dispositional predictors of organizational citizenship behaviors. *Personnel Psychology*, 48, 775-802.
- Ostroff, C. (1992). The relationship between satisfaction, attitudes and performance: An organizational level of analysis. *Journal of Applied Psychology*, 72, 963-974.
- Podsakoff, P.M., MacKenzie, S.B.; Lee, J.; Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903
- Schoorman, F.D., Mayer, R.C., & Davis, J.H. (1996, April). *Empowerment in veterinary clinics: The role of trust in delegation*. Paper presented at the 11th annual meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Schriesheim, C.A., & Gardiner, C.C. (1992). An exploration of the discriminant validity of the leader-member exchange scale (LMX7) commonly used in organizational research. In M, Schnake (Ed.). Proceedings of The Southern Management Association (pp, 91-93), Valdosta. GA: Southern Management Association.
- Schriesheim, C.A., Castro, S.L., & Coglisier, C.C. (1999). Leader-member exchange research (LMX): A comprehensive review of theory, measurement, and data analytic practices. *The Leadership Quarterly*, 10, 63–113.

- Simonson, I., & Nye, P. (1992). The effect of accountability on susceptibility to decision errors. *Organizational Behavior and Human Decision Processes*, 51, 416-446.
- Spector, P. E. (2006). Method variance in organizational research: truth or urban legend?. *Organizational Research Methods*, 9, 221–32.
- Tessor, A., & Rosen, S. (1975). The reluctance to transmit bad news. In L Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 193-232). New York: Academic Press.
- Tetlock, P. E. (1985). Accountability: The neglected social context of judgment and choice. *Research in Organizational Behavior*, 7, 297-332.
- Thomas, S.L. & Bretz, R.D. (1994). Research and practice in performance appraisal: Evaluating employee performance in America's largest companies. *S.A.M. Advanced Management Journal*, 59, 28-34.

Appendix A
Student Survey

PIN#: _____

What is your gender?

Male ____

Female ____

What is your race?

White ____

White, non-Hispanic ____

African-American ____

Hispanic ____

Asian-Pacific Islander ____

Native American ____

Other ____

What is your age?

What is the highest level of education you have completed?

4 Year College Degree (BS, BA) ____ Master's Degree ____ PhD ____

Trust

Please rate item 1 through 17 on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*)

1	2	3	4	5
<i>Strongly Disagree</i>	<i>Disagree</i>	<i>Undecided</i>	<i>Agree</i>	<i>Strongly Agree</i>

1. My adviser is very capable of performing his or her job
2. My adviser is known to be successful at the things he or she tries to do
3. My adviser has much knowledge about the work that needs done
4. I feel very confident about my adviser's skills
5. My adviser has specialized capabilities that can increase our performance
6. My adviser is well qualified
7. My adviser is very concerned about my welfare
8. My needs and desires are very important to my adviser
9. My adviser would not knowingly do anything to hurt me
10. My adviser really looks out for what is important to me
11. My adviser will go out of his or her way to help me
12. My adviser tried hard to be fair in dealings with others
13. My adviser has a strong sense of justice
14. I never have to wonder whether my adviser will stick to his or her word
15. My adviser's actions and behaviors are not very consistent
16. I like my adviser's values
17. Sound principles seem to guide my adviser's behavior

Leader-Member Exchange-Multidimensional

Please rate items 18 through 29 on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*)

1	2	3	4	5
<i>Strongly Disagree</i>	<i>Disagree</i>	<i>Undecided</i>	<i>Agree</i>	<i>Strongly Agree</i>

18. I do work for my adviser that goes beyond what is specified
19. I am willing to apply extra efforts, beyond those normally required, to meet my adviser's work goals.
20. I do not mind working my hardest for my adviser.
21. I like my adviser very much as a person
22. My adviser is the kind of person one would like to have as a friend.
23. My adviser is a lot of fun to work with.
24. My adviser defends my work actions to a superior, even without complete knowledge of the issue in question.
25. My adviser would come to my defense if I were 'attacked' by others.
26. My adviser would defend me to others if I made an honest mistake
27. I am impressed with my adviser's knowledge of his/her job.
28. I respect my adviser's knowledge of and competence on the job.
29. I admire my adviser's professional skills.

Frequency of Observation

Please rate item number 30 on a scale of 1 (*Not Sufficient*) to 5 (*Very Sufficient*)

1	2	3	4	
5				
<i>Not Sufficient</i> <i>Very Sufficient</i>	<i>Slightly Sufficient</i>	<i>Moderately Sufficient</i>	<i>Sufficient</i>	

30. How sufficient was the information gathered by your adviser about your performance in order to evaluate you?

Please rate item number 31 on a scale of 1 (*None*) to 5 (*A Great Deal*)

1	2	3	4	5
<i>None</i> <i>Great Deal</i>	<i>Little</i>	<i>Somewhat</i>	<i>Much</i>	<i>A</i>

31. When you are observed, how much time does your adviser spend observing you?

Please rate item numbers 32 and 33 on a scale of 1 (*Never*) to 5 (*Frequently*)

1	2	3	4	5
<i>Never</i> <i>Frequently</i>	<i>Very Rarely</i>	<i>Rarely</i>	<i>Occasionally</i>	

Appendix B
Adviser Survey

PIN#: _____

What is your gender?

Male ___

Female ___

What is your race?

White ___

White, non-Hispanic ___

African-American ___

Hispanic ___

Asian-Pacific Islander ___

Native American ___

Other ___

What is your age?

Are you a tenured professor?

Yes___ No___

Leader-Member Exchange-Multidimensional

Please rate each item on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*)

1	2	3	4	5
<i>Strongly Disagree</i>	<i>Disagree</i>	<i>Undecided</i>	<i>Agree</i>	<i>Strongly Agree</i>

1. I provide support and resources for my student that goes beyond what is specified
2. I am willing to apply extra efforts, beyond those normally required, to help my student meet goals
3. I do not mind working my hardest for my student.
4. I like my student very much as a person.
5. My student is the kind of person one would like to have as a friend.
6. My student is a lot of fun to work with.
7. My student defends my decisions, even without complete knowledge of the issue in question
8. My student would come to my defense if I were 'attacked' by others.
9. My student would defend me to others if I made an honest mistake
10. I am impressed with my student's knowledge of his/her field.
11. I respect my student's knowledge of and competence
12. I admire my student's professional skills.

Frequency of Observation

Please rate item number 13 on a scale of 1 (*Not Sufficient*) to 5 (*Very Sufficient*)

1	2	3	4	5
<i>Not Sufficient</i>	<i>Slightly Sufficient</i>	<i>Moderately Sufficient</i>		
<i>Sufficient</i>	<i>Very Sufficient</i>			

13. How sufficient was the information you gathered on your student's performance in order to evaluate them?

Please rate item number 14 on a scale of 1 (*Never*) to 5 (*Frequently*)

1	2	3	4	5
<i>Never</i>	<i>Very Rarely</i>	<i>Rarely</i>	<i>Occasionally</i>	
<i>Frequently</i>				

14. How frequently do you observe your advisee?
 15. How often do you meet with our advisee?

Please rate item number 16 on a scale of 1 (*None*) to 5 (*A Great Deal*)

1	2	3	4	5
<i>None</i>	<i>Little</i>	<i>Somewhat</i>	<i>Much</i>	
<i>A Great Deal</i>				

16. When you observe, how much time do you spend observing?

Accountability

Please rate item 17 through 20 on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*)

1	2	3	4	5
<i>Strongly Disagree</i>	<i>Disagree</i>	<i>Undecided</i>	<i>Agree</i>	<i>Strongly Agree</i>

17. I feel accountable to my advisee for the completeness of my performance evaluation
18. The completeness of my performance evaluations is linked to my relationship with my advisee
19. I believe my department chair has read the evaluation I made
20. I believe my colleagues care about the evaluation I made

Performance

Please rate your student on a 1 (*Not at All*) to 4 (*A Great Extent*) scale

1	2	3	4
<i>Not at All</i>	<i>Very Little</i>	<i>Somewhat</i>	<i>A Great Extent</i>

21. How **dependable** is your student in his/her graduate student roll?

Please rate your student on a 1 (*Very Poor*) to 5 (*Very Good*) scale

1	2	3	4	5
<i>Very Poor</i>	<i>Poor</i>	<i>Barely Acceptable</i>	<i>Good</i>	<i>Very Good</i>

22. What is the overall **quality** of work of your graduate student?

Please rate your student on a 1 (*Never*) to 5 (*Frequently*) scale

1	2	3	4	5
<i>Never</i>	<i>Very Rarely</i>	<i>Rarely</i>	<i>Occasionally</i>	<i>Frequently</i>

23. How often does your graduate student participate in **organizational citizenship behaviors** (e.g., helpful to others, stays longer to complete a task, does not complain, considers others, offers suggestions for change)?

Please rate your student on a 1 (*Never*) to 5 (*Always*) scale

1	2	3	4	5
---	---	---	---	---

*Never**Rarely**Sometimes**Often**Always*

- 24.** How **timely** has your student been in reaching graduate school milestones (e.g. completing course work, proposing or defending thesis/dissertation, taking comprehensive exam).

Appendix C

Adviser Recruitment Letter

We are currently conducting a study on graduate students' perceptions of the performance evaluation and feedback system in their department. The principal investigators of this project are Anne Scaduto, Penn State University, and Kevin Murphy, Penn State University. This is dissertation research so your help is greatly appreciated!

We would like you to participate in this interesting study. The requirements for participants are the following: (a) **forward the attached student recruitment email with student survey link to ONE of your advisees** so we can ask him or her to complete an online survey (completion will take about 15 minutes), and (b) **complete an online survey yourself** (this should take about 10 minutes). We ask you to forward on the recruitment email to your student because we want to keep student email addresses confidential.

As you know teaching and research are very important for University's success. Therefore, we would greatly appreciate your participation in this important study. If you are willing to participate please click on the link below to review the official implied consent form and complete the survey. Your **PIN number** is in the original email. Your advisee will have an identical PIN number so we can match your data. If you would prefer not to participate, you can simply delete this email.

Thank you for considering our request. If you have concerns or questions about this study, please contact Anne Scaduto (azs105@psu.edu).

Appendix D

Student Recruitment Letter

We are currently conducting a study on graduate students' perceptions of the performance evaluation and feedback system in their department. The principal investigators of this project are Anne Scaduto, Penn State University, and Kevin Murphy, Penn State University. This is dissertation research so your help is greatly appreciated!

We would like you to participate in this interesting study in conjunction with your adviser. The requirement for participants is the following: (a) **complete an online survey** (this should take about 15 minutes) keeping your adviser in mind as a referent.

As you know teaching and research are very important for University success. Therefore, we would greatly appreciate your participation in this important study. If you are willing to participate please click on the link below to review the official implied consent form and complete the survey. Your **PIN number** is in the original email sent to your adviser. Your adviser will have an identical PIN number so we can match your data. If you would prefer not to participate, you can simply delete this email.

Thank you for considering our request. If you have concerns or questions about this study, please contact Anne Scaduto (azs105@psu.edu).

Appendix E

Moderation Analyses: Student Ratings

The first facet of performance tested was ratings of organizational citizenship behaviors. A non-significant model emerged ($F(6, 89) = 1.81, n.s., R^2 = .11, \text{Adjusted } R^2 = .05$; see Table 13). Student ratings of LMX was a significant predictor of ratings of performance for this facet ($\beta = .50, p < .01$). The performance appraisal system characteristic variable was not found to be a significant moderator. The R^2 value before the cross products were entered was .06 and after the cross products were entered R^2 was .11; therefore change in R^2 was .05.

Table 13

Moderation Using Student Ratings: OCBs as the Dependent Variable

<i>Variable</i>	B	β	R^2	ΔR^2
Step 1			0.05	
Trust	-0.03	-0.3		
LMX	0.05	.02*		
Step 2			0.06	0.01
PA Sys Char	0.00	-0.1		
Step 3			0.08	0.02
Sys*LMX	0.00	0.42		
Sys*Trust	0.00	-0.41		
Step 4			0.11	0.03
Sys*LMX*Trust	0.00	-0.42		

Note: $R = .33, R^2 = .11$ and $\text{Adj. } R^2 = .05$

Note: $N = 95$

* $p < .05$

The second facet of performance tested was ratings of dependability. A non-significant model emerged ($F(6, 89) = 1.21, n.s., R^2 = .08, \text{Adjusted } R^2 = .01$; see Table

14). Student ratings of trustworthiness was a significant predictor of ratings of performance for this facet ($\beta = -.43, p < .05$). The performance appraisal system characteristic variable was not found to be a significant moderator. The R^2 value before the cross products were entered was .02 and after the cross products were entered R^2 was .08; therefore change in R^2 was .06.

Table 14

Moderation Using Student Ratings: Dependability as the Dependent Variable

<i>Variable</i>	B	β	R^2	ΔR^2
Step 1			0.01	
Trust	0	-0.13		
LMX	0.01	0.06		
Step 2			0.02	0.01
PA Sys Char	0.01	0.04		
Step 3			0.07	0.05
Sys*LMX	0.00	0.25		
Sys*Trust	0.00	-0.52		
Step 4			0.08	0.01
Sys*LMX*Trust	0.00	0.14		

Note: $R = .28, R^2 = .08$ and $Adj. R^2 = .01$

Note: $N = 95$

The third facet of performance tested was ratings of quality of work. A non-significant model emerged ($F(6, 89) = .74, n.s., R^2 = .05, Adjusted R^2 = -.02$; see Table 15). Neither student ratings of trustworthiness nor student ratings of LMX were found to be a significant predictor of ratings of performance for this facet. The performance appraisal system characteristic variable was not found to be a significant moderator. The

R^2 value before the cross products were entered was .02 and after the cross products were entered R^2 was .05; therefore change in R^2 was .03.

Table 15

Moderation Using Student Ratings: Quality of Work as the Dependent Variable

<i>Variable</i>	B	β	R^2	ΔR^2
Step 1			0.00	
Trust	0.00	-0.03		
LMX	0.01	0.06		
Step 2			0.02	0.02
PA Sys Char	0.01	0.17		
Step 3			0.05	0.03
Sys*LMX	0.00	0.32		
Sys*Trust	0.00	-0.17		
Step 4			0.05	0.03
Sys*LMX*Trust	0.00	-0.01		

Note: $R = .22$, $R^2 = .05$ and $\text{Adj. } R^2 = -.02$

Note: $N = 95$

The fourth facet of performance tested was ratings of timeliness. A non-significant model emerged ($F(6, 89) = .42$, *n.s.*, $R^2 = .03$, Adjusted $R^2 = -.04$; see Table 16). Neither student ratings of trustworthiness nor student ratings of LMX were found to be a significant predictor of ratings of performance for this facet. The performance appraisal system characteristic variable was not found to be a significant moderator. The R^2 value before the cross products were entered was .01 and after the cross products were entered R^2 was .03; therefore change in R^2 was .02.

Table 16

Moderation Using Student Ratings: Timeliness as the Dependent Variable

<i>Variable</i>	B	β	R^2	ΔR^2
Step 1			0.00	
Trust	0	-0.04		
LMX	0.01	0.08		
Step 2			0.01	0.01
PA Sys Char	0	0.07		
Step 3			0.01	0.00
Sys*LMX	0.00	0.19		
Sys*Trust	0.00	-0.14		
Step 4			0.03	0.02
Sys*LMX*Trust	0.00	-0.32		

Note: $R = .17$, $R^2 = .03$ and $\text{Adj. } R^2 = -.04$

Note: $N = 95$

Lastly, overall performance ratings were tested. A non-significant model emerged ($F(6, 89) = 1.03$, *n.s.*, $R^2 = .07$, Adjusted $R^2 = .00$; see Table 17). Neither student ratings of trustworthiness nor student ratings of LMX were found to be a significant predictor of ratings of performance for this facet. The performance appraisal system characteristic variable was not found to be a significant moderator. The R^2 value before the cross products were entered was .03 and after the cross products were entered R^2 was .07; therefore change in R^2 was .04.

Table 17

Moderation Using Student Ratings: Overall Performance as the Dependent Variable

<i>Variable</i>	B	β	R²	ΔR^2
Step 1			0.02	
Trust	-0.05	-0.19		
LMX	0.08	0.22		
Step 2			0.03	0.01
PA Sys Char	0.02	0.11		
Step 3			0.06	0.03
Sys*LMX	0.01	0.46		
Sys*Trust	0.00	-0.47		
Step 4			0.07	0.01
Sys*LMX*Trust	0.00	-0.23		

Note: R = .26, R² = .07 and Adj. R² = .00

Note: N = 95

Appendix F

Moderation Analyses: Adviser Ratings

The first facet of performance tested was ratings of organizational citizenship behaviors. A significant model emerged ($F(6, 89) = 5.81, p < .001, R^2 = .28, \text{Adjusted } R^2 = .23$; see Table 18). Adviser ratings of LMX was a significant predictor of ratings of performance for this facet ($\beta = .36, p < .001$). Student ratings of trust was also a significant predictor of OCBs ($\beta = .27, p < .05$). The performance appraisal system characteristic variable was found to be a significant moderator ($\beta = .23, p < .05$). The R^2 value before the cross products were entered was .19 and after the cross products were entered R^2 was .28; therefore change in R^2 was .09.

Table 18

Moderation Using Adviser Ratings: OCBs as the Dependent Variable

<i>Variable</i>	B	β	R^2	ΔR^2
Step 1			0.19	
Trust	0.01	0.13		
LMX	0.06	.41**		
Step 2			0.19	0.00
PA Sys Char	0.00	-0.01		
Step 3			0.26	0.07
Sys*LMX	-0.01	-0.17		
Sys*Trust	0.01	.27*		
Step 4			0.28	0.02
Sys*LMX*Trust	0.00	-0.17		

Note: $R = .53, R^2 = .28$ and $\text{Adj. } R^2 = .23$

Note: $N = 95$

* $p < .05, **p < .001$

The second facet of performance tested was ratings of dependability. A significant model emerged ($F(6, 89) = 4.34, p < .01, R^2 = .23, \text{Adjusted } R^2 = .17$; see Table 19).

Adviser ratings of LMX was a significant predictor of ratings of performance for this facet ($\beta = .42, p < .001$). The performance appraisal system characteristic variable was not found to be a significant moderator. The R^2 value before the cross products were entered was .20 and after the cross products were entered R^2 was .23; therefore change in R^2 was .03.

Table 19

Moderation Using Adviser Ratings: Dependability as the Dependent Variable

<i>Variable</i>	<i>B</i>	β	R^2	ΔR^2
Step 1			0.19	
Trust	0	0.05		
LMX	0.02	.44*		
Step 2			0.20	0.01
PA Sys Char	0.01	0.10		
Step 3			0.22	0.02
Sys*LMX	0.00	0.12		
Sys*Trust	0.00	0.12		
Step 4			0.23	0.01
Sys*LMX*Trust	0.00	0.07		

Note: $R = .48, R^2 = .23$ and $\text{Adj. } R^2 = .17$

Note: $N = 95$

* $p < .05$

The third facet of performance tested was ratings of quality of work. A significant model emerged ($F(6, 89) = 7.36, p < .001, R^2 = .33, \text{Adjusted } R^2 = .29$; see Table 20).

Adviser ratings of LMX were found to be a significant predictor of ratings of performance for this facet ($\beta = .51, p < .001$). The performance appraisal system

characteristic variable was not found to be a significant moderator. The R^2 value before the cross products were entered was .30 and after the cross products were entered R^2 was .33; therefore change in R^2 was .03.

Table 20

Moderation Using Adviser Ratings: Quality of Work as the Dependent Variable

<i>Variable</i>	<i>B</i>	β	R^2	ΔR^2
Step 1			0.30	
Trust	0.00	0.04		
LMX	0.07	.54*		
Step 2			0.30	0.00
PA Sys Char	0.01	0.03		
Step 3			0.33	0.03
Sys*LMX	0.00	-0.12		
Sys*Trust	0.00	0.19		
Step 4			0.33	0.00
Sys*LMX*Trust	0.00	-0.05		

Note: $R = .57$, $R^2 = .33$ and $\text{Adj. } R^2 = .29$

Note: $N = 95$

* $p < .001$

The fourth facet of performance tested was ratings of timeliness. A non-significant model emerged ($F(6, 89) = 2.10$, *n.s.*, $R^2 = .12$, Adjusted $R^2 = -.07$; see Table 21). Neither student ratings of trustworthiness nor adviser ratings of LMX were found to be a significant predictor of ratings of performance for this facet. The performance appraisal system characteristic variable was not found to be a significant moderator. The R^2 value before the cross products were entered was .08 and after the cross products were entered R^2 was .12; therefore change in R^2 was .04.

Table 21

Moderation Using Adviser Ratings: Timeliness as the Dependent Variable

<i>Variable</i>	B	β	R^2	ΔR^2
Step 1			0.05	
Trust	-0.01	-0.1		
LMX	0.03	.20*		
Step 2			0.08	0.03
PA Sys Char	0.04	0.19		
Step 3			0.12	0.04
Sys*LMX	0.00	0.07		
Sys*Trust	0.00	0.21		
Step 4			0.12	0.00
Sys*LMX*Trust	0.00	-0.01		

Note: $R = .35$, $R^2 = .12$ and $\text{Adj. } R^2 = -.07$

Note: $N = 95$

* $p < .05$

Lastly, overall performance ratings were tested. A significant model emerged ($F(6, 89) = 7.64, p < .001, R^2 = .07, \text{Adjusted } R^2 = .00$; see Table 22). Adviser ratings of LMX were found to be a significant predictor of ratings of performance for this facet ($\beta = .46, p < .001$). Additionally, the performance appraisal system characteristic variable was found to be a significant moderator ($\beta = .24, p < .05$). The R^2 value before the cross products were entered was .03 and after the cross products were entered R^2 was .07; therefore change in R^2 was .04.

Table 22

Moderation Using Adviser Ratings: Overall Performance as the Dependent Variable

<i>Variable</i>	B	β	R²	ΔR^2
Step 1			0.02	
Trust	-0.05	-0.19		
LMX	0.08	0.22		
Step 2			0.03	0.01
PA Sys Char	0.02	0.11		
Step 3			0.06	0.03
Sys*LMX	0.01	0.46		
Sys*Trust	0.00	-0.47		
Step 4			0.07	0.01
Sys*LMX*Trust	0.00	-0.23		

Note: R = .26, R² = .07 and Adj. R² = .00

Note: N = 95

Appendix G

Mediation Analyses: Student Ratings

The Baron and Kenny (1986) method was used to test the mediation hypotheses at the individual level of analysis (students). The first mediation tested was student ratings of overall performance mediating the relationship between student ratings of LMX and the feedback reaction satisfaction. The relationship between LMX and satisfaction, LMX and student ratings of overall performance, and student ratings of overall performance and satisfaction are $\beta = .59, p < .001$, $\beta = .07, n.s.$, and $\beta = .13, n.s.$ respectively. Based on these non-significant results, Step 4 was not completed. This indicated that student ratings of overall performance do not mediate the relationship between LMX and satisfaction. The relationship between LMX and satisfaction had an R^2 of .35. See Table 23 for R^2 values.

The second mediation tested was student ratings of overall performance mediating the relationship between student ratings of LMX and the feedback reaction motivation. The relationship between LMX and motivation, LMX and student ratings of overall performance, and student ratings of overall performance and motivation are $\beta = .58, p < .001$, $\beta = .07, n.s.$, and $\beta = .10, p < .05$ respectively. Based on the significant result in the second relationship, Step 4 was completed. The simple regression for Step 4 resulted in a continued significant relationship between LMX and motivation ($\beta = .56, p < .001$) and the mediator was significantly related to motivation ($\beta = .20, p < .05$). The relationship between LMX and motivation had an R^2 of .33. When the mediator was included, the R^2 value was .37 making change in R^2 .04. This could indicate that student ratings of overall

performance partially mediate the relationship between LMX and motivation. Table 24 displays R^2 values.

The third mediation tested was student ratings of overall performance mediating the relationship between student ratings of LMX and the feedback reaction accuracy. The relationship between LMX and accuracy, LMX and student ratings of overall performance, and student ratings of overall performance and accuracy are $\beta = .14, n.s.$, $\beta = .07, n.s.$, and $\beta = .17, n.s.$ respectively. Based on these non-significant results, Step 4 was not completed. This indicated that student ratings of overall performance do not mediate the relationship between LMX and accuracy. The relationship between LMX and accuracy had an R^2 of .02.

The fourth mediation tested was student ratings of overall performance mediating the relationship between trustworthiness and the feedback reaction satisfaction. The relationship between trustworthiness and satisfaction, trustworthiness and student ratings of overall performance, and student ratings of overall performance and satisfaction are $\beta = .63, p < .001$, $\beta = -.02, n.s.$, and $\beta = .13, n.s.$ respectively. Based on these non-significant results, Step 4 was not completed. This indicated that student ratings of overall performance do not mediate the relationship between trustworthiness and satisfaction. The relationship between trustworthiness and satisfaction had an R^2 of .40. See Table 25 for R^2 values.

The fifth mediation tested was student ratings of overall performance mediating the relationship between trustworthiness and the feedback reaction motivation. The relationship between trustworthiness and motivation, trustworthiness and student ratings of overall performance, and student ratings of overall performance and motivation are $\beta =$

.47, $p < .001$, $\beta = -.02$, *n.s.*, and $\beta = .24$, $p < .05$ respectively. Based on these significant results, Step 4 was completed. The simple regression for Step 4 resulted in a continued significant relationship between trustworthiness and motivation ($\beta = .25$, $p < .01$) and the mediator was significantly related to motivation ($\beta = .47$, $p < .001$). The relationship between trustworthiness and motivation had an R^2 of .22. When the mediator was included, the R^2 value was .28 making change in R^2 .06. This could indicate that student ratings of overall performance partially mediate the relationship between trustworthiness and motivation. Table 26 displays the R^2 values.

The sixth mediation tested was student ratings of overall performance mediating the relationship between trustworthiness and the feedback reaction accuracy. The relationship between trustworthiness and accuracy, trustworthiness and student ratings of overall performance, and student ratings of overall performance and accuracy are $\beta = .14$, *n.s.*, $\beta = -.02$, *n.s.*, and $\beta = .19$, *n.s.* respectively. Based on these non-significant results, Step 4 was not completed. This indicated that student ratings of overall performance do not mediate the relationship between trustworthiness and accuracy. The relationship between trustworthiness and accuracy had an R^2 of .02.

Due to the low power of the Baron and Kenny (1986) technique possibly explaining the failure to find mediation, the researcher compared the R^2 values of the mediator (student ratings of overall performance) regressed onto Y (feedback reactions), X (student ratings of LMX or trustworthiness) regressed onto Y, and X and M regressed onto Y. This method could not be used to test if student ratings of overall performance mediate the relationship between LMX and accuracy because neither LMX nor student ratings of overall performance significantly related to this feedback reaction. The

researcher also could not use this approach when testing if student ratings of overall performance mediate the relationship between trustworthiness and accuracy because neither trustworthiness nor student ratings of overall performance significantly related to accuracy. Tables 23 through 26 present the relevant R^2 values.

Table 23

Student Ratings of Performance Mediating the Relationship between Student ratings of LMX and Satisfaction

<i>Variable</i>	R^2
LMX	0.35
Performance Ratings	0.02
LMX + Performance Ratings	0.35

Table 24

Student Ratings of Performance Mediating the Relationship between Student ratings of LMX and Motivation

<i>Variable</i>	<i>R²</i>
LMX	0.33
Performance Ratings	0.06
LMX + Performance Ratings	0.37

Table 25

Student Ratings of Performance Mediating the Relationship between Trustworthiness and Satisfaction

<i>Variable</i>	<i>R²</i>
Trustworthiness	0.40
Performance Ratings	0.02
Trustworthiness + Performance Ratings	0.42

Table 26

Student Ratings of Performance Mediating the Relationship between Trustworthiness and Motivation

<i>Variable</i>	<i>R²</i>
Trustworthiness	0.22
Performance Ratings	0.06
Trustworthiness + Performance Ratings	0.29

Based on these comparisons, it appears that performance ratings explain additional variance in the relationship between trustworthiness and motivation. The results displayed in Tables 23 through 26 make it evident that there is no strong mediator effect that would be found with more powerful techniques. These results further suggest that there is no advantage to considering alternative mediation approaches to the Baron and Kenny (1986) method. Although there is some evidence of partial mediation, this may not mean more than the fact that X (LMX or trustworthiness), Y (feedback reactions), and M (student ratings of overall performance) are all positively correlated; therefore this aspect of the model is not supported.

Appendix H

Mediation Analyses: Adviser Ratings

The Baron and Kenny (1986) method was used to test the mediation hypotheses at the individual level of analysis (advisers). The first mediation tested was adviser ratings of overall performance mediating the relationship between adviser ratings of LMX and the feedback reaction satisfaction. The relationship between LMX and satisfaction, LMX and adviser ratings of overall performance, and adviser ratings of overall performance and satisfaction are $\beta = .18$, *n.s.*, $\beta = .52$, $p < .001$, and $\beta = .19$, *n.s.* respectively. Based on these non-significant results, Step 4 was not completed. This indicated that adviser ratings of overall performance do not mediate the relationship between LMX and satisfaction. The relationship between LMX and satisfaction had an R^2 of .03.

The second mediation tested was adviser ratings of overall performance mediating the relationship between adviser ratings of LMX and the feedback reaction motivation. The relationship between LMX and motivation, LMX and adviser ratings of overall performance, and adviser ratings of overall performance and motivation are $\beta = .04$, *n.s.*, $\beta = .52$, $p < .001$, and $\beta = .10$, *n.s.* respectively. Based on these non-significant results, Step 4 was not completed. This indicated that adviser ratings of overall performance do not mediate the relationship between LMX and motivation. The relationship between LMX and motivation had an R^2 of .00.

The third mediation tested was adviser ratings of overall performance mediating the relationship between adviser ratings of LMX and the feedback reaction accuracy. The relationship between LMX and accuracy, LMX and adviser ratings of overall performance, and adviser ratings of overall performance and accuracy are $\beta = .02$, *n.s.*, $\beta = .52$, $p < .001$, and $\beta = .07$, *n.s.* respectively. Based on these non-significant results, Step 4 was not completed. This

indicated that adviser ratings of overall performance do not mediate the relationship between LMX and accuracy. The relationship between LMX and accuracy had an R^2 of .00.

The fourth mediation tested was adviser ratings of overall performance mediating the relationship between trustworthiness and the feedback reaction satisfaction. The relationship between trustworthiness and satisfaction, trustworthiness and adviser ratings of overall performance, and adviser ratings of overall performance and satisfaction are $\beta = .62, p < .001$, $\beta = .05, p < .001$, and $\beta = .19, n.s.$ respectively. Based on the non-significant result in the third relationship, Step 4 was not completed. This indicated that adviser ratings of overall performance do not mediate the relationship between trustworthiness and satisfaction. The relationship between trustworthiness and satisfaction had an R^2 of .40. See Table 27 for R^2 values.

The fifth mediation tested was adviser ratings of overall performance mediating the relationship between trustworthiness and the feedback reaction motivation. The relationship between trustworthiness and motivation, trustworthiness and adviser ratings of overall performance, and adviser ratings of overall performance and motivation are $\beta = .47, p < .001$, $\beta = .05, p < .001$, and $\beta = .10, n.s.$ respectively. Based on the non-significant result in the third relationship, Step 4 was not completed. This indicated that adviser ratings of overall performance do not mediate the relationship between trustworthiness and motivation. The relationship between trustworthiness and motivation had an R^2 of .22. Table 28 displays R^2 values.

The sixth mediation tested was adviser ratings of overall performance mediating the relationship between trustworthiness and the feedback reaction accuracy. The relationship between trustworthiness and accuracy, trustworthiness and adviser ratings of overall performance, and adviser ratings of overall performance and accuracy are $\beta = .14, n.s.$, $\beta = .05, p < .001$, and $\beta = .07, n.s.$ respectively. Based on these non-significant results, Step 4 was not completed. This indicated that adviser ratings of overall performance do not mediate the relationship between

trustworthiness and accuracy. The relationship between trustworthiness and accuracy had an R^2 of .02.

Due to the low power of the Baron and Kenny (1986) technique possibly explaining the failure to find mediation, the researcher compared the R^2 values of the mediator (adviser ratings of overall performance) regressed onto Y (feedback reactions), X (adviser ratings of LMX or trustworthiness) regressed onto Y, and X and M regressed onto Y. This method could not be used to test if adviser ratings of overall performance mediate the relationship between trustworthiness and any of the feedback reactions because neither trustworthiness nor adviser ratings of overall performance significantly related to any of the feedback reactions. The researcher also could not use this approach when testing if adviser ratings of overall performance mediate the relationship between trustworthiness and accuracy because neither trustworthiness nor adviser ratings of overall performance were significantly related to accuracy. Tables 27 and 28 present the relevant R^2 values.

Table 27

Adviser Ratings of Performance Mediating the Relationship between Trustworthiness and Satisfaction

<i>Variable</i>	R^2
Trustworthiness	0.40
Performance Ratings	0.04
Trustworthiness + Performance Ratings	0.42

Table 28

Adviser Ratings of Performance Mediating the Relationship between Trustworthiness and Motivation

<i>Variable</i>	R^2
Trustworthiness	0.22
Performance Ratings	0.01
Trustworthiness + Performance Ratings	0.22

Based on these comparisons, it appears that performance ratings do not explain much additional variance. The results displayed in Tables 27 through 28 make it evident that there is no strong mediator effect that would be found with more powerful techniques.

These results further suggest that there is no advantage to considering alternative mediation approaches to the Baron and Kenny (1986) method.

VITA

Anne Scaduto

ACADEMIC PREPARATION

The Pennsylvania State University, University Park, PA

August 2006-May 2011

- Ph.D. in Industrial/Organizational Psychology

Old Dominion University, Norfolk, VA

August 2004-December 2006

- M.S. in Psychology

The Pennsylvania State University, University Park, PA

August 2000-May 2004

- B.S. in Psychology

- Minor in Spanish, Minor in Business

CONSULTING EXPERIENCE

Human Capital Consultant, Federal Management Partners, Alexandria, VA, Present

Intern, Valtera Corporation, Houston, TX, 2008-2009

RESEARCH EXPERIENCE

Research Assistant, Center for the Study of Terrorism, 2009-2010

Research Assistant, Marine Corps Project, 2007

COLLEGE TEACHING EXPERIENCE

Teaching Assistant, Psychology: Graduate Statistics, 2007-2008

Teaching Assistant, Psychology: Introduction to Organizational Psychology, Introduction to Psychology and Introduction to Developmental Psychology, 2006-2007

PUBLICATIONS

Scaduto, A., Lindsay, D., & Chiaburu, D.S. (2008). Leader influences on training transfer and intervening mechanisms. *International Journal of Training and Development*, 12(3), 158-170.