

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600

UMI<sup>®</sup>



The selection of raters in a multi-source feedback process

Sandra Petosa

A Thesis

In

The John Molson

School Of Business

Presented in Partial Fulfillment of the Requirements  
For the Degree of Master of Science in Administration at  
Concordia University  
Montreal, Quebec, Canada

April 2001

©Sandra Petosa, 2001

**CONCORDIA UNIVERSITY**



National Library  
of Canada

Acquisitions and  
Bibliographic Services

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque nationale  
du Canada

Acquisitions et  
services bibliographiques

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-59289-8

Canada

## Abstract

### The selection of raters in a multi-source feedback process

Sandra Petosa

In recent years, multi-source feedback (MSF) has emerged as a popular performance appraisal method. This method promises performance evaluation data that is less biased, more reliable, and more valid than the traditional supervisor-only appraisal method. However, given that it is common practice in a MSF process for ratees to select the raters of their performance, the method may actually be introducing new biases. Given this, the current study employed a web-based survey to assess the effects that an individual's level of organizational based self-esteem (OBSE) may have on rater selection. The hypothesis that individuals low in OBSE would be more likely than individuals high in OBSE to select raters that were perceived as being good friends was supported by some of the analyses. In addition, the hypothesis that individuals high in OBSE would be more likely than individuals low in OBSE to choose raters that were perceived as being highly acquainted with their work was also supported by some of the analyses.

The assumption behind these hypotheses is that choosing friends as raters will lead to positively enhanced and unrealistic assessments while choosing raters that are highly task acquainted will lead to accurate and thereby potentially negative assessments. While critical performance assessments may not be desired in the short-term, the long-term growth and developmental consequences are substantial. In addition, the possible effects that the rater source (peer/subordinate) and the purpose of the appraisal (administrative/developmental) may have on the process were also examined.

## **Dedication**

To my loving parents.

## Table of Contents

<b>List of Figures.....</b>	<b>vii</b>
<b>List of Tables .....</b>	<b>viii</b>
<b>INTRODUCTION.....</b>	<b>1</b>
<b>Feedback.....</b>	<b>1</b>
The definition of feedback .....	1
The importance of accurate feedback.....	1
<b>Performance Appraisal.....</b>	<b>3</b>
Rater biases .....	4
Reducing rater biases .....	7
<b>Introduction of MSF .....</b>	<b>9</b>
MSF and the reduction of rater biases.....	10
MSF and the creation of new biases.....	12
<b>Active versus Passive Feedback Seekers .....</b>	<b>14</b>
<b>Impression Management: Ingratiation .....</b>	<b>17</b>
<b>Past Research On Rater Selection .....</b>	<b>20</b>
<b>Current study .....</b>	<b>22</b>
<b>Organization-Based Self-Esteem .....</b>	<b>25</b>
<b>Self-Evaluation .....</b>	<b>30</b>
Theory of Self-Enhancement.....	31
Theory of Self-Consistency.....	34
Theory of Self-Assessment .....	35
Enhancement, Consistency or Assessment?.....	36
Cognitive .....	39
Affective.....	40
<b>Friendship.....</b>	<b>43</b>
<b>Task Acquaintanceship.....</b>	<b>46</b>
<b>Administrative versus Developmental .....</b>	<b>48</b>
<b>Subordinates/Peers.....</b>	<b>52</b>
<b>METHOD .....</b>	<b>56</b>
<b>Design .....</b>	<b>56</b>
<b>Sample.....</b>	<b>56</b>
<b>Experimental Manipulation .....</b>	<b>58</b>
<b>Pilot Test .....</b>	<b>59</b>
<b>Ethical Provisions.....</b>	<b>60</b>
<b>Web-Based Survey .....</b>	<b>60</b>
<b>Measures .....</b>	<b>63</b>
Demographic variables .....	63
Organizational based self-esteem.....	63
Assessments for each peer and subordinate .....	64
Friendship.....	65
Task acquaintanceship.....	66
Selection of raters.....	67
<b>ANALYSIS .....</b>	<b>68</b>
<b>Descriptive Analyses .....</b>	<b>68</b>
<b>Main Analyses: Tests of the hypotheses .....</b>	<b>71</b>
Friendship.....	74
Task Acquaintanceship .....	78

Length of time .....	81
<b>Main Analyses: Tests of the hypotheses (Individual Data).....</b>	<b>82</b>
Friendship.....	86
Task Acquaintanceship .....	91
Length of Time.....	96
<b>DISCUSSION.....</b>	<b>98</b>
Review of Results.....	98
Hypotheses .....	100
Limitations .....	105
Directions for future research .....	106
<b>CONCLUSION .....</b>	<b>109</b>
<b>References.....</b>	<b>110</b>
<b>APPENDIX 1.....</b>	<b>122</b>
<b>APPENDIX 2.....</b>	<b>123</b>
<b>APPENDIX 3.....</b>	<b>124</b>



## List of Figures

<b>Diagram 1: Variables involved in the rater selection process.....</b>	<b>24</b>
<b>Figure 1: OBSE Frequency .....</b>	<b>70</b>
<b>Figure 2: Source by OBSE interaction for friendship .....</b>	<b>77</b>
<b>Figure 3: Selection by purpose interaction: Friendship (Individual data) .....</b>	<b>88</b>
<b>Figure 4: Selection by OBSE interaction for friendship: Administrative condition (Individual data).....</b>	<b>90</b>
<b>Figure 5: Selection by purpose interaction: Task acquaintanceship (Individual data) .....</b>	<b>94</b>
<b>Figure 6: Selection by OBSE interaction for task acquaintanceship: Developmental condition (Individual data).....</b>	<b>95</b>

## List of Tables

<b>Table 1: Means, standard deviations and correlations (N = 91)</b>	<b>68-69</b>
<b>Table 2: Friendship Variables</b>	<b>71</b>
<b>Table 3: Task Acquaintanceship Variables</b>	<b>72</b>
<b>Table 4: Mean friendship score for each condition</b>	<b>73</b>
<b>Table 5: Mean task acquaintanceship score for each condition</b>	<b>73</b>
<b>Table 6: RM results for the administrative condition: friendship outcome</b>	<b>75</b>
<b>Table 7: RM results for the developmental condition: friendship outcome</b>	<b>75</b>
<b>Table 8: RM results for the administrative condition: task acquaintanceship outcome</b>	<b>79</b>
<b>Table 9: RM results for the developmental condition: task acquaintanceship outcome</b>	<b>79</b>
<b>Table 10: Length of Time results for the administrative condition</b>	<b>82</b>
<b>Table 11: Length of Time results for the developmental condition</b>	<b>82</b>
<b>Table 12: Friendship Variables (Individual Data)</b>	<b>83</b>
<b>Table 13: Task Acquaintanceship Variables (Individual Data)</b>	<b>84</b>
<b>Table 14: Mean friendship score for each condition (Individual Data)</b>	<b>85</b>
<b>Table 15: Mean task acquaintanceship score for each condition (Individual Data)</b>	<b>85</b>
<b>Table 16: RM results for the administrative condition: friendship outcome (Individual Data)</b>	<b>86</b>
<b>Table 17: RM results for the developmental condition: friendship outcome (Individual Data)</b>	<b>86</b>
<b>Table 18: RM results for the administrative condition: task acquaintanceship outcome (Individual Data)</b>	<b>92</b>
<b>Table 19: RM results for the developmental condition: task acquaintanceship outcome (Individual Data)</b>	<b>92</b>
<b>Table 20: Length of Time results for the administrative condition (Individual Data)</b>	<b>97</b>
<b>Table 21: Length of Time results for the developmental condition (Individual Data)</b>	<b>97</b>

# **INTRODUCTION**

## **Feedback**

### **The definition of feedback**

Ashford and Cummings (1983) define feedback as a subset of information available to individuals in their environment that denotes how well they are meeting various goals. In addition, Herold and Geller (1977) state that feedback includes both information about the appropriate behaviors to achieve a goal (referent information) and information about how well an individual in enacting those behaviors (appraisal information). Therefore, feedback forms an important resource and as a result, much research has been directed towards understanding the variables that allow organizations to make the best use of it.

### **The importance of accurate feedback**

The term accuracy will be used throughout this paper to denote the extent to which the evaluations provided by the rater approximates the target employee's objective and actual performance. Therefore, accurate performance evaluations inevitably includes a description of both positive and negative behaviors performed at work. This definition implies that there exists a "true performance" score, which is essentially impossible to obtain. Nonetheless, the goal of accuracy is implicit in every appraisal process (DeNisi, Cafferty & Meglino, 1984). The impact that feedback has on

the enhancement of both performance and motivation at work has been an accepted psychological principle since the early 1950s (Payne & Haunty, 1955). Accurate self-assessments have instrumental value and have been associated with enhanced individual and organizational outcomes (Yammarino & Atwater, 1993). For example, accurate self-evaluations have been associated with promotions whereas inflated self-evaluations have been associated with career derailment (McCall & Lombardo, 1983). Accurate performance feedback plays a significant role in achieving goals, as well as motivating and developing employees (Ammons, 1956; Chapanis, 1964; Ilgen, Fisher & Taylor, 1979). It has even been suggested that without accurate performance feedback organizational deterioration is inevitable (Taylor, Fisher & Ilgen, 1984).

Given that feedback about job performance is evaluative information, it is either positive or negative, it is inherently affective (Ashford & Cummings, 1983). Given this affective quality, it has been shown that individuals do not behave rationally when seeking feedback (Ashford & Tsui, 1991). Individuals may seek positive feedback intentionally in order to avoid a negative evaluation of their behavior at work. Alternatively, individuals may avoid accurate and therefore potentially negative performance information because it threatens their ego (Miller, 1976) or because it makes their negative qualities much more apparent to both the rater and the other organizational members who are exposed to the feedback. However, avoiding negative performance information in the work setting can be very costly. Employees need to know what they are doing wrong and the areas of their work that need improvement if they are to grow and develop themselves (Ashford & Tsui, 1991). If employees are not

made aware of their weaknesses, then they cannot take the steps to correct their behavior. The fact that individuals often hold inaccurate and over-inflated views of their own behavior accentuates this problem (see Mabe & West, 1982 for a review). Individuals often believe that they are performing better than they actually are at work. Due to this, it is especially important for individuals to seek feedback from raters who will rate them accurately. This would allow their self-assessments to become more realistic.

One study conducted by Ashford and Tsui (1991) found that those employees who did seek out negative feedback had a more tempered view of their abilities, their performance, and their standing in the organization. They found that seeking negative feedback was positively associated with having an accurate perception of how others evaluate them, whereas the tendency to seek positive feedback was not associated with such accuracy. Employees with an accurate view of how their constituents were responding to and evaluating their work allowed them to better meet their demands and keep their efforts in line with their goals. In order to be effective and prosper on the job, an accurate view of performance is necessary.

### **Performance Appraisal**

The most typical and common example of a formal feedback generating process within an organization is the performance appraisal. Performance appraisals refer to the process by which individual performance is assessed for organizational purposes.

Typically, this process takes the form of an observer, often the immediate supervisor, rating the job performance of an employee. The individual who does the rating is often referred to as the rater and the individual being rated is often referred to as the ratee or target. Performance appraisals began in the late 1800s as the industrial age mechanized production processes and increased individual output became mandatory. Initially, supervisors solely commented on employees, without using a set or formalized evaluation form. However, in the 1960s and the 1970s, the management by objectives (MBO) approach which compared individual performance to organizational goals and strategies became popular and forms and scales eventually evolved. In most organizations these evaluations are conducted annually or semi-annually and a bias-free evaluation is essential for accuracy to be maintained.

### Rater biases

Traditionally, the sole rater of employee performance was the immediate supervisor. The immediate supervisor was judged to be in the best position to evaluate and monitor each subordinate. However, the number one reason why traditional performance appraisals have difficulty yielding accurate results is because the manager as the sole rater often lacks sufficient information concerning the employee's actual performance (Longnecker & Goff, 1990). The second reason why traditional performance appraisals fall short is an extension of the first. Basically, managers do not have sufficient information to accurately evaluate an employee's behavior. As a result, they rely on impressions and self-selected recollections, leading to errors and biases

(Longnecker & Goff, 1990). Rating errors reduce reliability, validity and the ultimate utility of the appraisal since inaccurate results are gathered (Roberts, 1998). Reliability can be defined as "the extent to which a set of measurements is free from random-error-variance"(Guion, 1965, p.30). Validity on the other hand can be defined as the extent to which an instrument measures what it is purported to measure (Guion, 1980). Therefore, examining the biases that may operate in any performance appraisal system warrants attention.

Some of the most common errors that occur with traditional performance appraisals are: 1) the halo error, 2) the similar-to-me error, 3) the central tendency error, 4) the leniency or strictness error, 5) the contrast error and 6) the recency error (Hedge & Kavanagh, 1988). These errors and biases are all individual rater errors. They are errors and biases that occur either consciously or unconsciously within the rater.

The halo or horns error is the most common type of error in traditional performance appraisals (Belcourt, Sherman, Bohlander & Snell, 1996). The halo error occurs when one distinct incident of desirable behavior overshadows all subsequent behaviors. The evaluator distinctly recalls one incidence in which the employee did something well and then selectively ignores all negative information pertaining to that employee, noticing and remembering only the good (Balzer & Sulsky, 1992).

The central tendency error is one in which the rater evaluates all ratees as average. The rater does not differentiate between high and low performers and just

lumps all ratees into the middle of the scale. This is especially common when one individual is responsible for evaluating several employees. In contrast to the central tendency error, it is also common for some raters to give unusually high or low ratings. The leniency or strictness error is one in which the rater evaluates all employees as either very good or very poor. For example, a manager may erroneously assert “all my employee are excellent” or conversely “none of my employee are good enough”. These errors often occur when the ratee is sure to find out how they were rated by their manager. There is no anonymity in the results and therefore supervisors are often hesitant to rate honestly. For example, one study revealed that supervisors, as the sole raters, often do not evaluate subordinates accurately so that they can avoid a negative situation (Fried, Tiegs & Bellamy, 1992).

The contrast error occurs when an employee’s evaluation is biased either positively or negatively because of a comparison with the performance of another employee (Belcourt et al., 1996). For example, an average employee may appear especially productive when compared to a poor performer. However, the same employee may appear unproductive when compared to a star performer. Contrast errors are likely to occur under the forced distribution method in which employees are directly compared to one another (Murphy & Cleveland, 1991).

Some rating errors are temporal in nature. For example, when the appraisal is based largely on the employee’s recent behavior, good or bad, the rater has committed the recency error (Gacalone, 1989). In one survey, employees stated that if a suspension



occurred at the beginning of the year, it was often not discussed in the performance evaluation at the end of the year. However, if it occurred in the two months before the appraisal, it was the main focus of the evaluation (Roberts, 1998). Lastly, the similar-to-me error occurs when raters inflate the evaluations of people with whom they have something in common. The bias lies in the fact that the ratee is not being evaluated on his or her behaviors at work but rather on their similarity and likeability to the rater.

### Reducing rater biases

The rating errors just described are very real and many organizations have attempted to reduce them either by influencing the rating format or by training raters about the potential biases that may exist (DeNisi et al., 1984). For example, one procedure to reduce the leniency or strictness error is to clearly define the characteristics or dimensions of performance and to provide descriptions of behaviors, known as anchors on a scale (Murphy & Cleveland, 1991). Another approach is to require ratings to conform to a forced distribution. Managers appraising employees under a forced distribution system would be required to place a certain percentage of employees into various performance categories (Murphy & Cleveland, 1991). In addition, the recency error can be reduced by training the rater to routinely document employee accomplishments or failures throughout the appraisal period (Bernardin & Beatty, 1984).

However, many of the rater errors and biases that were just reviewed occur because only one individual is responsible for evaluating a given employee's performance (Westerman & Rosse, 1997). Although these errors would not be completely abolished if multiple raters were used, they would be minimized. The existence of multiple raters reduces the idiosyncrasies and biases inherent in single-source-assessments and allows for the reporting of different perspectives (Bernardin & Beatty, 1984; Cascio, 1987; Cederblom & Lounsbury, 1980). Although traditionally the immediate supervisor was judged to be the best rater of employee performance, it becomes clear upon reflection that the immediate supervisor is often not in the best position to judge employee behavior. Frequently, supervisors are too far removed from their subordinates or are too busy performing other tasks, such as attending meetings or preparing budget reports. In addition, the increasing complexity of organizational life points to the value of receiving input from people who have different experiences with and perceptions of the target individual (Becker & Klimoski, 1989). Organizations are becoming flatter and self-managed work teams are increasing, making the traditional supervisory performance appraisal unsuitable. Jobs are becoming increasingly complex and it is often unrealistic to assume that one person can fully and accurately rate an employee's performance (Westerman & Rosse, 1997).

As a result, the process of providing individuals with feedback from several sources, including coworkers, subordinates, clients, and supervisors has emerged as a very popular performance evaluation technique. Such a process has often been referred to as: multi-rater feedback, multi-source feedback, full-circle evaluations and 360-

degree feedback. All terms generally refer to the concept of having several individuals evaluate the target employee's work performance, but the difference lies in how many raters are involved in the evaluation. Multi-rater and multi-source feedback refers to the process by which performance evaluations are collected from more than one source, for example by peers, subordinates and the immediate supervisor (London & Smither, 1995). Full-circle evaluations or 360-degree feedback on the other hand refers more specifically to the process of having a full circle of individuals (self, peers, subordinates, supervisors, customers, suppliers, etc.) evaluate the target employee (London & Smither, 1995). Since this paper will focus on the evaluations provided by peers and subordinates and not the whole circle of potential raters, the term multi-source feedback (MSF) will be used throughout the remainder of this paper.

### **Introduction of MSF**

Although MSF was relatively unknown in the business world twenty years ago, today it is being introduced into 90% of Fortune 1000 companies and is continuing to spread among smaller firms as well (Coates, 1998; Atwater & Waldman, 1998). MSF processes are usually administered once or twice a year and the evaluations are typically gathered through ratings and written verbatim comments on a set of items assessing job-relevant dimensions or competencies (Dalessio, 1998). The feedback is then compiled into a report that more often than not protects the anonymity of the raters. Given this anonymity of the data, the ratings are less likely to be inflated (Antonioni, 1994). Most often the ratings from different sources are explicitly contrasted on the same feedback

page (London & Smither, 1995). Therefore, the ratings provided by peers would be averaged and compared to the ratings provided by subordinates for example. This gives the ratee different and possibly inconsistent evaluations of the different performance dimensions assessed.

In addition, unlike the traditional supervisor-only appraisal, MSF ratings are often intended to serve a developmental purpose (London & Smither, 1995). For example, Brutus and Derayeh (2000), in their random sample of large organizations, found that 74% of the companies interviewed used MSF solely for developmental purposes. This means that the performance evaluations remained private and belonged to the ratee for personal development. Nonetheless, Brutus and Derayeh (2000) also found that 52% of the companies that had these so-called development programs also used them in one way or another to influence administrative decisions, such as compensation, promotions or lateral transfers. In addition, 23% of the companies interviewed used the MSF evaluations for both developmental and administrative purposes (Brutus & Derayeh, 2000). Therefore, it is clear that both purposes are important and active in business practice.

#### MSF and the reduction of rater biases

One of the reasons why MSF has emerged as a very popular performance evaluation technique is because more valid and reliable data is likely to be gathered. Since the biases that each rater may possess are idiosyncratic, they average out amongst

the raters (Coates, 1998). It is a basic psychometric principle that the greater the number of sources evaluating someone or something, the less bias and greater validity you get. Consider the halo or horns error and the recency error. If multiple raters were involved in the evaluation, the biases would average out amongst the raters. Multiple raters would provide an enhanced opportunity to observe and measure job behaviors therefore one behavior would not dominate the evaluation (Borman, 1974). For example, one study examined the differential effects of leniency and halo errors for supervisor only ratings versus supervisor, self and peer ratings. They found that the multiple rating scores were less biased than the single rating scores (Holzbach, 1978).

Using multiple rating sources leads to increased reliability, fairness and observational power since raters with differing perspectives and role relationships with the ratee evaluate a target employee (Harris & Scharbroeck, 1988; Latham & Wexley, 1982; Borman, 1974). Multi-source feedback promises each employee a more democratic and perhaps less biased view of how to improve performance and it eases the supervisor's burden of being the sole judge of performance (Bohl, 1996). One survey conducted in over 750 companies revealed that multi-source feedback was superior to supervisor-only appraisals in encouraging communication between the manager and the employee and in promoting positive change in work behaviors (Bohl, 1996).

The utility of the MSF process rests primarily on the validity of the performance information collected from the various raters. The assumption is that MSF processes will help ratees better understand how others view their work performance and therefore

they should develop a more accurate and reliable sense of their performance and goal accomplishment (Tornow, 1993). Rater errors and biases are reduced and the evaluation data is much more useful. However, the current paper will evaluate the validity of the added benefits that a MSF provides when ratees are given the freedom to choose who, among their many peers and subordinates, will evaluate their performance. In a MSF process, ratees are often given the freedom to choose which peers and subordinates they would like to rate their behavior. For example, Brutus and Derayeh (2000) found that 84% of the companies interviewed gave their employees complete freedom in choosing their raters. Therefore, it is very possible that through this practice new biases may be manifesting themselves.

#### MSF and the creation of new biases

In a MSF process, the target individual is often instructed to personally select and distribute all of their feedback questionnaires to a mixed group of peers and subordinates (Dalessio, 1998; Sederburg & Rogelberg, 1998). This commonly used form of distribution is intended to increase rater motivation since they have personal contact with the ratee (Church, Rogelberg & Wacławski, 2000). In addition, feedback acceptance and perceived fairness of the process is enhanced when ratees are allowed to participate in the design and implementation of the process (Gilliland & Langdon, 1998). This increased acceptance of the process results in a greater motivation to change undesirable behaviors (Farr & Newman, 2000).

Nonetheless, the potential for personal motives to operate during this distribution process is substantial. Ratees may choose raters in a manner that is self-enhancing in order to reap the benefits of a higher performance rating. As stated by Herold and Parsons (1985): “multi-source feedback may be the result of a proactive individual seeking, interpreting, and otherwise generating his or her own feedback” (p.290). If this were the case, the MSF process would no longer be generating accurate, valid and useful information about employee performance. As a result, the performance appraisal data would not serve its intended purpose of developing employees or acting as a sound basis upon which administrative decisions can be made (Farr & Newman, 2000).

In summary, we have seen so far that in order for feedback to be beneficial, it must be accurate and free from biases. Without accurate feedback information, improvements in performance behavior is unlikely to occur. In response to this, most organizations have abandoned the traditional supervisor only performance evaluation method and have adopted multi-source feedback. This process allows for more accurate and less biased feedback to be gathered since several perspectives are tapped into. However, given that during a MSF process ratees are allowed to choose their raters, new biases may be manifesting themselves. This time however, the biases may be primarily *ratee* biases rather than *rater* biases.

As of yet, very little research has examined the impact that the ratee may have in influencing feedback results although we have seen that 1) a substantial literature on rater biases does exist and 2) the practical importance of this issue is pressing due to the

phenomenal growth of MSF processes that allow ratees to select raters. In addition, there is a theoretical readiness for examining the potential for ratee biases. This theoretical readiness comes from Ashford and Cummings' (1983) argument that feedback is not solely an organizational resource but a personal resource as well. Personal goals (the attainment of a raise or promotion) as well as organizational goals (increased results) may be in operation when seeking feedback. This shift points to the fact that the ratee is an *active* seeker and manipulator of his/her feedback environment and supports the idea that ratees may select raters in a manner that allows them to influence their evaluation results.

### **Active versus Passive Feedback Seekers**

Although feedback is an important organizational resource, Ashford and Cummings (1983) in their now classic article argue that this exclusive focus on feedback as an organizational resource has constrained our complete understanding of the feedback process. They argue that the feedback-seeking process is a complex one and that *individual* goals such as maintaining a positive self-image or receiving a promotion often interact with *organizational* goals of obtaining accurate feedback to improve employee performance at work. This suggests that the selection of raters may be influenced by the ratee's personal agenda (Farr & Newman, 2000).

Ashford and Cummings (1983) list some ways in which feedback may act as a personal resource. Individuals may use feedback seeking to create an enhanced view of



their performance at work or as an ego defense mechanism. Individuals may seek out positive feedback in order to defend themselves against the damage that a negative evaluation may have. For example, at the time of a performance evaluation, employees may actively seek feedback from certain choice individuals whom they believe will rate them positively in an effort to boost their ratings and consequently their self-esteem and positive self-image at work.

The fact that Ashford and Cummings (1983) argue that feedback may be a personal resource as well as an organizational one is especially interesting because it points to the fact that individuals are active in shaping their feedback environments. Originally, the dominant view surrounding feedback exchange portrayed the image of a *sender* conveying some sort of message to the *receiver* (Ilgen et al., 1979). The feedback model was rather simplistic and did not consider the role that the recipient of the feedback may play in the process. The message was portrayed as flowing from the source to the recipient in a uni-directional manner with the receiver acting passively (Herold & Fedor, 1998). These traditional models of feedback are often referred to as cybernetic models (Carver & Scheier, 1981). These models compare the feedback process to a thermostat, with behavior change resulting from the delivery of feedback that compares performance to some desired or standard level.

However, Ashford and Cummings (1983) proposed that individuals do not simply wait around for feedback. Instead, they actively seek it and shape their feedback environments accordingly. The two processes in which individuals actively seek out

feedback have been labeled monitoring and inquiry (Ashford & Cummings, 1983). Therefore, individuals either actively monitor their information environment for personally relevant information or they directly inquire the information from others around them.

Monitoring involves attending to the environment, observing the behaviors and cues that are present and useful as feedback. For example, when an employee observes who is invited to certain meetings or takes note of the management behaviors that are rewarded. Monitoring is therefore subtle and involves a great amount of interpretation and inference. Since the feedback cues are a function of the individual's self-related schema, it is very possible that individuals may interpret the situation very differently from the way it is in actuality (Ashford & Cummings, 1983). Inquiry on the other hand is much more straightforward and involves directly asking individuals in the environment for their perception and evaluation of the behavior in question. For example, directly asking your boss if you did a good job on your last project.

In addition, Fedor (1991) has proposed a new model of the feedback process that traces how recipients receive and respond to the feedback message about their performance. The recipients of the feedback are seen as the "central processors" and are portrayed as active gathers of the performance information. Recipients' characteristics or individual differences, such as their level of self-esteem, are expected to influence all the phases of the feedback process, namely, their perception of the source, their

perception of the message characteristics and behavioral intention once the feedback is absorbed.

Therefore, we have witnessed a shift in the ideology surrounding feedback-seeking behavior. Feedback recipients are active agents and individual differences are likely to impact upon how the message is perceived and reacted to. As a result, recipients have the potential to use organizational feedback as a personal resource. Research interested in studying the feedback environment can no longer simply take organizational goals into account. Personal goals and motives of the feedback seeker must also be considered. As a result, a large body of research has examined the social influence processes that are ongoing in organizations. These social influences all refer to behaviors that are performed by organizational incumbents to manipulate or influence other organizational members or the situation. This research will be reviewed in the next section.

### **Impression Management: Ingratiation**

The most notable social influence processes that have been studied are upward influence, organizational politics and impression management (Liden & Mitchell, 1988). Upward influence has been defined as “attempts to influence someone higher in the formal hierarchy of authority within the organization” and is often directed towards immediate supervisors (Porter, Allen & Angle, 1983, p.409). In contrast, organizational politics has been defined by Mayes and Allen (1977, p.675) as “using influence to

obtain ends not sanctioned by the organization”. Lastly, impression management involves manipulating the impression others hold of you in an organizational setting to make sure you are seen in a positive manner (Schlenker, 1980). The importance and relevance of these phenomena is that they all point to the fact that organizational members are not passive agents in their environments. Rather, they are active in managing the impressions that others hold of them and in influencing the activities around them.

Leary and Kowalski (1990) proposed a model of impression management composed of two processes: *impression motivation* and *impression construction*. Impression motivation refers to the incentives that drive people to manage the impressions others hold of them. These incentives include: 1) social and material outcomes, 2) self-esteem maintenance and 3) the development of an identity. The first incentive involves conveying a desirable image in order to maximize rewards and avoid punishments or undesirable outcomes. The second incentive refers to a drive to maintain a favorable personal image so that self-esteem is maintained or raised and not deflated. Lastly, the incentive to create a positive personal identity may also drive impression management behaviors.

Impression construction on the other hand refers to the behaviors that are enacted in order to manage impressions. Therefore, in a MSF process, individuals may construct the impressions others hold of them by choosing raters who will rate them positively. Similarly, they may be motivated to do this in order to maximize the rewards that may

follow a favorable evaluation, to maintain their level of self-esteem or to form a positive personal identity.

In addition, impression management strategies can be divided into two categories: defensive and assertive (Fedor, 1991). Defensive strategies would be those most likely to operate in response to negative feedback. They are often used to deflect blame and involve refusing responsibility for the negative evaluation, providing excuses, apologies and the expression of remorse (Fedor, 1991). Assertive strategies on the other hand operate most in reaction to positive evaluations and involve self-promotion and ingratiation tactics that emphasize the recipient's positive qualities (Fedor, 1991).

As a result, ingratiating behaviors are often defined as one particular example of impression management strategies. Ingratiation has been defined as “a set of assertive tactics which have the purpose of gaining the approbation of an audience that controls significant awards for the actor” (Tedeschi & Melburg, 1984, p.37). Therefore, ingratiation involves actively attempting to present oneself in an enhanced or positive manner to another. Liden and Mitchell (1988) propose a model of ingratiating behavior that details the causes of ingratiation, the manner in which individuals may ingratiate, and the effects that this behavior has on the target, the individual, and the organization. These authors classify ingratiation strategies into three categories: self-presentation, target-directed, and third-person directed (Liden and Mitchell, 1988). The self-presentation ingratiation strategy is one in which an individual creates a positive image of him or herself in general. With the target-directed strategy, the individual attempts to

make the target feel good so that a favorable impression of the ingratiation is created. Lastly, in the third-person ingratiation strategy, the ingratiation attempts to portray a positive impression of him or herself to a third party in the hopes that this impression will then in turn be relayed back to the target (Liden and Mitchell, 1988).

Therefore, the ingratiation literature is one of the few lines of research that has examined the influence that an organizational member may have in influencing the opinion that the target holds of their work performance. Consequently, this line of research can be considered a stepping stone for examining the potential influencing effects that a ratee may have in a MSF process that gives the focal individual the opportunity to select their own raters. However, does there exist any other research that has more directly examined this issue?

### **Past Research On Rater Selection**

To date, the formal selection of feedback information has been acknowledged by the users of MSF but rarely investigated empirically. As a result, the recommendations surrounding the issue of rater selection are often inconsistent. For example, Dalessio (1998) suggests that the ratee should be active in choosing raters in order to promote empowerment and acceptance of the results. In contrast, Van Velsor, Leslie, and Fleenor, (1997) suggest that the raters should be chosen by the ratee's manager. Nonetheless, Farr and Newman (2000) provide a good starting point by discussing the issues and implications surrounding rater selection in a MSF process. They discuss the

importance of ensuring the collection of accurate and valid data through a MSF that allows the ratee to choose his/her raters. Nonetheless, no empirical evidence was gathered in this article.

Westerman & Rosse (1997) and Church et al. (2000) touched upon the issue of rater selection when they examined whether or not selected raters took the time to complete the evaluation. They explored the variables that may play a role in determining rater response rates. Given that most MSF processes are both anonymous and voluntary, some raters may choose not to participate in the process. Westerman & Rosse (1997) found that those raters who did participate felt more confident regarding the rating format, did not feel strong pressure to bias their feedback, felt a high degree of confidence in the accuracy of their ratings, felt secure in their rating role, and were likely to perceive that the ratings would influence outcomes. In addition, Church et al. (2000) explored the relationship between rater response rates and performance. They found that the number of rater responses that ratees received was unrelated to their level of performance as assessed through a measure of service quality.

Stepanski and Fisicaro (1999) are the only authors to explore the effects of having ratees choose their raters in a MSF process. Specifically, they examined the effects of having managers select their own raters versus obtaining ratings from all possible sources. They hypothesized that when ratees were allowed to choose their raters, the resultant evaluations would be higher than when they were not allowed to choose raters.

In this study, the 360-degree feedback ratings of mid-level managers in a medium-sized Midwestern U.S. company were utilized. The feedback ratings were obtained from five types of raters (self, supervisor, clients, immediate peers and subordinates) for two consecutive years. Self and supervisor ratings were the same in year one and two and ratees chose their client raters in both years. However, ratees chose their peer and subordinate raters only in year one, not in year two. The analysis of the data revealed that self, supervisor and client ratings between year one and year two were the same. No difference between the year one client ratings and the year two client ratings were obtained. However, peer and subordinate ratings were higher in year one, when they were given the opportunity to select their raters than in year two when they were not given the opportunity to select their raters. Therefore, Stepanski and Fisicaro (1999) found that self-selecting raters led to enhanced performance evaluations. This study is therefore a good starting point for examining the impact that rater selection may have. However, why these differences may exist, the underlying mechanisms or strategies that individuals may use to select raters, as well as the individual and situational variables that may impact upon the process have yet to be explored.

### **Current study**

The current study will set out to directly explore the mechanisms that are involved when a ratee is asked to select performance evaluation raters. For example, if Mary has ten peers and ten subordinates and is asked to distribute five of her evaluation forms to her peers and five to her subordinates, who will she choose? Which five out of



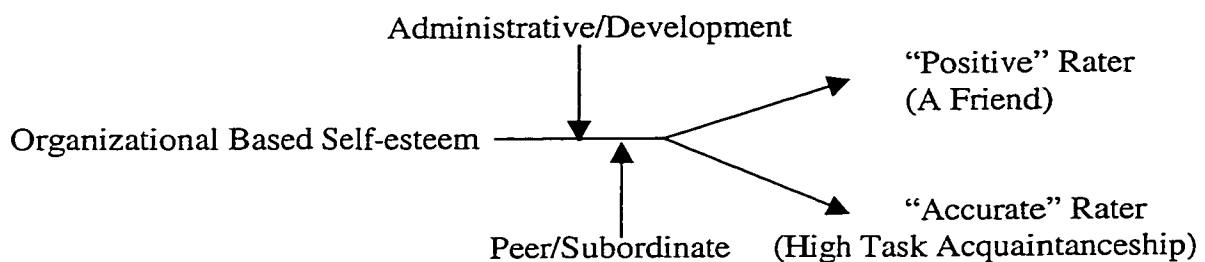
her ten peers will she decide to give her evaluation forms to? Conversely, which five out of her ten subordinates will she decide to give her evaluation forms to? In addition, will Mary's level of self-esteem influence whether she will choose someone who she knows will rate her "positively" versus someone whom she believes is more likely to give an "accurate" assessment? Individuals low in self-esteem may be more likely to behave opportunistically and attempt to propagate positive impressions of themselves, whereas individuals high in self-esteem may simply seek realistic appraisals of their performance, so they can grow and develop.

As a result, individuals seeking "positive" assessments may be more likely to choose those peers and subordinates that they like and are friends with. In contrast, individuals seeking "accurate" assessments may be more likely to choose those peers and subordinates that are high on task acquaintanceship or know a lot about their day-to-day work.

In addition, what situational variables are likely to influence whether "positive" or "accurate" raters are chosen? Will it make a difference if the performance evaluation information is to be used for administrative versus developmental purposes? When evaluation information is intended to be used for administrative purposes, the results are available to the ratee's supervisor and could influence decisions related to compensation, promotion, terminations or lateral transfers. Conversely, when evaluation information is intended to be used solely for developmental purposes, no one but the ratee, not even his or her immediate supervisor would see the results and the

evaluations would be used for personal development only. One can imagine that ratees may be more likely to choose “positive” raters when the evaluation is intended to be used for administrative purposes because it can directly impact upon their rewards and the image that others hold concerning their level of performance at work.

Lastly, because of the different power dynamics inherent in peer versus subordinate relationships, different strategies may operate when ratees are choosing peer raters versus subordinate raters. The variables that will be explored in this study are summarized in Diagram 1 below.



**Diagram 1: Variables involved in the rater selection process**

As of yet, there is no empirical evidence supporting the existence of these strategies even though the implications are substantial. The findings of this study may be useful for several groups: raters, ratees, practitioners delivering feedback in various settings, business consultants and individuals in administrative positions who might be designing or implementing such systems in organizations.

## **Organization-Based Self-Esteem**

One of the most significant dimensions along which employees differ is that of self-esteem. Self-esteem is widely recognized as one of the most influential determinants of behavior, cognition and affect (Brockner, 1988). Given this, the level of ratee self-esteem may very well play a role in determining which raters are chosen in a MSF process. It may not be the case that all individuals will select raters who will rate them positively in an effort to enhance their evaluations. If this were the case, then all evaluations would be enhanced and the bias would not be as important since it would apply consistently to all assessments. However, since OBSE may interact with selection to make some individuals select raters who will evaluate them positively and others to select raters who will provide more accurate assessments, some evaluations may be enhanced and inaccurate and others may not. This makes the potential for bias in a MSF process worth investigating.

The concept of self-esteem is especially relevant when discussing performance appraisals because feedback by definition aims to alter one's sense of self. The relationship between performance feedback and self-esteem at work is a circular one. For example, if a positive performance evaluation is received this will add to the employee's personal esteem that he/she is a good worker. The fact that he/she is a good worker will then in turn influence the type of performance feedback received. Performance feedback can therefore either enhance or reduce one's concept of self and social persona (Northcraft & Ashford, 1990).

Self-esteem is the evaluative component of a broader representation of self, the self concept (Blascovich & Tomaka, 1991; Wylie, 1974). The construct of self-esteem is often used synonymously with such terms as “self-acceptance”, “self-worth”, “self-confidence”, “self-assurance”, and “self-efficacy” (Brockner, 1988). However, all differ slightly from self-esteem. For example, self-acceptance refers to an individual’s attitudes toward their self-esteem, and self-worth to an individual’s perceptions of their value to either themselves or others (Brockner, 1988). Self-confidence and self-assurance on the other hand seem to relate to Bandura’s (1977) concept of self-efficacy, which refers to an individual’s belief that they have the ability to successfully execute some behavior. Although there exist several definitions of self-esteem, most agree that it refers to the degree to which an individual likes or dislikes oneself (Brockner, 1988). It is a personal evaluation reflecting what people think of themselves as individuals, the degree to which they positively or negatively evaluate themselves and believe themselves to be capable and worthy (Coopersmith, 1967; Gelfand, 1962; Korman, 1976; and Wells & Marwell, 1976).

Self-esteem can influence work behaviors in two ways (Brockner, 1988). First, employees bring to the job different levels of self-esteem, which in turn influences how they think, feel and behave at work. Second, since people generally need to feel good about themselves, they routinely engage in behaviors or thoughts that enhance, preserve, or restore their self-esteem. As a consequence, high self-esteem individuals differ greatly from low self-esteem individuals in the ways they think, feel and behave at work. For example, individuals high in self-esteem are less negatively affected by

chronic stressors, such as role ambiguity and conflict (Mossholder, Bedeian & Armenakis, 1981), less likely to imitate the managerial styles of their supervisors (Weiss, 1977) and more likely to want promotions for reasons of justice and enhanced responsibility rather than status (Locke, 1977).

Nonetheless, the concept of self-esteem is a hierarchical and multifaceted phenomenon (Shavelson, Hubner & Stanton, 1976; Song & Hattie, 1985; and Tharenou, 1979). Korman (1970) characterized self-esteem as encompassing three levels of generality, the first being global or chronic self-esteem. Global or chronic self-esteem refers to the relatively enduring perception of overall worth and competence that an individual has for his or her self. An individual low in chronic self-esteem, as a result of past failures and frustrations, possesses the self-image of a person who is generally inadequate and incapable of satisfying personal needs. On the other hand, the individual high on global self-esteem has the self-image of a person who is generally adequate and capable of needs satisfaction (Korman, 1966). However, in addition to chronic self-esteem, two situational components of the self-concept exist: task or situation specific self-esteem and social self-esteem. Task or situation specific self-esteem refers to a person's self-evaluation based on their behavior in a specific situation or task. Social self-esteem on the other hand refers to the self-evaluation that results from the expectations that others have of a person in a situation.

Based on this level of generality, Pierce, Gardner, Cummings, and Dunham (1989) developed the construct of organizational based self-esteem (OBSE). OBSE is a

person's self-evaluation of how he or she performs in a specific situation, work. They define organization-based self-esteem as:

"The degree to which organizational members believe that they can satisfy their needs by participating in roles within the context of an organization. People with high OBSE have a sense of personal adequacy as organizational members and a sense of having satisfied needs from their organizational roles in the past." (p. 625)

Therefore, OBSE is the extent to which organizational members feel that they are important, meaningful, effective, and worthwhile in the organization they work for. It is much more narrow and specific than global self-esteem although it has been found to be positively correlated with global self-esteem (Pierce et al., 1989). In addition, just like global self-esteem, OBSE is part of an individual's basic belief system and as such has been found to be stable over time as long as no major changes in the work environment occur (Pierce et al., 1989).

Theoretically, OBSE is intended to differ from such value-laden constructs as central life interest (Dubin, 1956) and job involvement (Lohdahl & Kejner, 1965), which possess higher emotional-affective components. OBSE is also intended to be different from such outcomes as self-perceptions of efficacy in performing a particular task (Bandura, 1977). OBSE reflects an individual's self-perceived competence within an organization whereas general self-efficacy reflects a belief that self-perceived competence can be translated into actions that will result in successful performance.

Some of the antecedents that have been positively correlated with OBSE are managerial respect, mechanistic organizational designs, and job complexity (Pierce et al., 1989). In addition, OBSE has been found to lead to increased intrinsic work motivation, job performance, job satisfaction, the engagement of organizationally beneficial behaviors, organizational commitment, and organizational satisfaction (Pierce et al., 1989). These results are very similar to other examinations of self-esteem in that positive experiences seem to lead to high self-esteem and negative experiences to low self-esteem.

The rationale and need for the construct of organizational based self-esteem is apparent when one considers a review conducted by Tharenou (1979). She discovered that while global self-esteem scales are likely to be more appropriate for studies of individuals within the context of total life events, situation specific measures are more appropriate for very specific behaviors. In addition, Epstein's (1979) research on the relationship between behaviors and attitudes suggests that the more self-esteem is framed within a context consistent with the behavior or attitude it seeks to predict, the higher will be the observed correlation between the variables. Therefore, organization-based self-esteem is more appropriate than global self-esteem in predicting any behaviors studied in the context of the work environment. However, do there exist any theories that can help to explain the mechanisms or processes of these work behaviors? Three self-evaluation theories that may help explain why individuals with differing levels of OBSE may choose different raters of their performance in a MSF process will be presented in the next section.

## **Self-Evaluation**

Individuals seek feedback and receive evaluations on a daily basis and in almost all aspects of their lives including home, work, school, leisure and play. These evaluations are a fundamental part of life and one of the most informative means of individual learning about the self-concept. The self-concept can be defined as the cognitive representation of one's own abilities (Kihlstrom & Cantor, 1984). However, these evaluations are motivated (Sedikides & Strube, 1997). Motives influence the ways in which people select self-relevant information, recall information, draw inferences about themselves and make future plans.

Three prominent theories that explain the possible motives that operate when dealing with self-evaluative information are the theory of self-enhancement, the theory of self-consistency and the theory of self-assessment. These theories may help explain the mechanisms that individuals with differing levels of OBSE may use when choosing raters in a MSF process. The fact that these motives exist has been demonstrated time and time again, however the conditions and individual differences that influence which motive operates at any given moment are less clear. Nonetheless, the usage of these theories in this study is worthwhile. These theories will help to explain all the variables discussed in this study thus far. Specifically, they will explain how different levels of self-esteem may operate to influence whether a positive rater versus an accurate and potentially negative rater of performance is chosen.



## Theory of Self-Enhancement

The theory of self-enhancement states that individuals seek to maximize their self-esteem and feelings of self-worth. According to the theory of self-enhancement, individuals actively protect their self-concepts from negative information and seek out positive information instead (Dipboye, 1977). A great deal of empirical evidence supports self-enhancement theory. For example, the motive to self-enhance has been used to interpret the recall and processing of self-relevant knowledge (Kuiper & Derry, 1982; Skowronski, Betz, Thompson, & Shannon, 1991), self-presentational strategies (Baumeister, 1982; Jones & Pittman, 1982; Schlenker, 1980), self-attributions (Bradley, 1978; Greenwald, 1980), predictions of future success (Alloy & Abramson, 1988; Taylor & Brown, 1988), and the targets to which people compare themselves (Taylor & Loebel, 1989; Tesser, 1986).

The theory proposes that all individuals have a need to view themselves as favorably as possible, resulting in an increase or maintenance of feelings of worth, competence, and satisfaction (simple self-enhancement). The theory also states that the strength of the need for self-enhancement is directly related to the extent to which this need had been thwarted in the past. Therefore, individuals low in self-esteem would be more likely to strive toward self-enhancement than their high self-esteem counterparts (compensatory self-enhancement). In addition, Brockner and Elkind (1985) found that high self-esteem individuals were more likely to exhibit a marked increase in work motivation and performance when confronted with negative feedback about their

performance whereas low self-esteem individuals demonstrated a sharp decline. This suggests that for individuals with a low level of self-esteem, a failure experience or a negative evaluation of their performance may diminish the effectiveness of their functioning, allowing themselves to become apathetic or to lose confidence in their abilities, including those that are not directly related to the feedback (Shrauger & Rosenberg, 1970). This would create a downward development slope for those individuals low in self-esteem. For individuals with high self-esteem however, it may stimulate them towards better achievement.

While individuals low in self-esteem should see negative performance information as valuable since it may help them correct errors and better attain their goals, they will also be highly motivated to maintain a favorable view of themselves and avoid this negative information (Brickman & Bulman, 1977; Jones & Gerard, 1967). For example, Larson (1989) stated that when employees suspect that they are poor performers, they will use feedback-seeking strategies that minimize the amount of negative performance feedback they receive. In addition, Northcraft and Ashford (1990) supported this thesis when they found that individuals with low performance expectations engaged in less feedback inquiry than those with high expectations. These employees sought evaluations that were image enhancing rather than diagnostic. As a result, they did not receive the information that they needed to improve and develop themselves at work (Morrison & Bies, 1991).

The long-term implications of this continual avoidance of negative feedback will be stunted individual growth or development. As was discussed at the start of this paper, negative feedback is essential in order to improve performance and keep individual behavior in line with organizational goals. If low self-esteem and thereby poor performing individuals are most likely to avoid negative feedback they will never be able to improve their performance. A vicious circle of thwarted growth will result.

Therefore, in a multi-source feedback process in which employees are given the opportunity to choose who will rate their performance, persons with a low level of organizational based self-esteem may be more likely to seek feedback that is positive in order to keep their level of motivation and performance from declining. Northcraft and Ashford (1990) found some preliminary support for this in their study that analyzed feedback-seeking behavior in a stock market simulation. The results revealed that high self-esteem participants requested more portfolio feedback than did low self-esteem participants. Perhaps the same relationship may hold in a MSF process.

In addition, Karl and Kopf (1994) examined this issue in their study when they assessed the impact of self-esteem and performance on students' willingness to seek videotaped feedback. They also found that those individuals who needed to improve their performance the most were the least likely to seek the feedback.

## Theory of Self-Consistency

Nonetheless, there exists a competing theoretical approach to self-enhancing theory, self-consistency theory (Korman, 1976). Self-consistency theory of work motivation emerged from the very popular cognitive consistency theories in the psychological literature, such as Festinger's (1957) dissonance theory, Heider's (1958) and Newcomb's (1959) balance theories, and Osgood and Tannenbaum's (1955) congruity theory. According to consistency theory, individuals are motivated to maintain a balance between their existing self-conceptions and new self-relevant information. Therefore, individuals with a high level of self-esteem should seek and respond more favorably to positive evaluations of themselves or their behavior whereas individuals with a low level of self-esteem should seek and respond more favorably to negative evaluations (Korman, 1976).

Several findings seem to support the theory of self-consistency. For example, individuals have been found to selectively recall information that supports rather than contradicts their self-beliefs (Swann & Read, 1981) and to discard self-refuting feedback as inaccurate (Frey & Stahlberg, 1986). Individuals have also been found to place less value on feedback sources that disconfirm their existing self-views (Frey, 1981; Shrauger & Lund, 1975).

Consistency theory would predict that in a multi-source feedback process, persons will attempt to choose raters that will give feedback consistent with their self-

views about their level of work performance. Specifically, individuals low in OBSE would be more likely to choose raters that will evaluate them accurately, and potentially negatively in order to maintain consistency with their personal self-evaluation. On the other hand, individuals high in OBSE will be more likely to choose raters that will evaluate them positively.

### Theory of Self-Assessment

The theory of self-assessment on the other hand states that individuals are inherently motivated to obtain an accurate evaluation of the self. Evaluation is undertaken in order to obtain a realistic appraisal of one's strengths and weaknesses (Troe, 1986; Weiner, 1980). Therefore, these individuals seek diagnostic information about themselves that is accurate, regardless of whether it is positive or negative or whether it is consistent with existing evaluations of the self. In order for the theory of self-assessment to operate effectively, there must be a relative degree of uncertainty about the attribute or performance dimension in question (Sedikides & Strube, 1997). If there is a high degree of certainty about the attribute being assessed then additional self-diagnostic information would be useless and the self-assessment motive will be diminished. This fits well with most performance appraisals at work. It is unlikely that any employee is completely certain about all of his or her work behaviors.

Most of the research that supports the theory of self-assessment illustrates that individuals will rate tasks that are highly diagnostic as more attractive than tasks that are

not very diagnostic (Strube & Roemmele, 1985). Furthermore, research seems to suggest that individuals high in self-esteem prefer engaging in high rather than low diagnostic tasks (Trope, 1975), and manifest greater willingness and stronger intentions to work on highly diagnostic tasks (Trope, 1979). Therefore, individuals high in OBSE may be more likely to operate under the self-assessment motive and seek out raters that will give an accurate evaluation of their work performance. These individuals may be more likely to take a chance at receiving negative information because they feel it will be instrumentally valuable and can be used to alter and improve their behavior (Northcraft & Ashford, 1990). In contrast, it is likely that individuals low in OBSE will feel threatened by the potential for negative feedback. Consequently, self-assessment theory would predict that in a multi-source feedback process, persons with a high level of self-esteem will be more likely than individuals low in self-esteem to choose raters that will evaluate them accurately.

#### Enhancement, Consistency or Assessment?

We have seen so far that all three motives may be adaptive and provide explanations for behavior under different circumstances. However, which motive would be most operative in a multi-source feedback process that allowed ratees to choose which peers and subordinates will rate their behavior? Given that the potential for negative feedback can be threatening to the self-concept, it is very likely that the ratee's level of OBSE will play a role in influencing which motive is operative.

Both the theory of self-enhancement and self-consistency predict that people with positive self-views work to maintain such views, albeit for different reasons. However, the two theories make competing predictions regarding individuals with low levels of self-esteem. Self-consistency theorists assume that individuals with low self-esteem prefer negative feedback because it is predictable and consistent with their self-evaluations. On the other hand, self-enhancement theorists assume that such individuals prefer positive feedback, more so than individuals with a high level of self-esteem, because they want to increase their feelings of worth. While several decades of research has tried to resolve this inconsistency, only mixed results have surfaced, with some favoring self-consistency theory and others favoring self-enhancement (Jones, 1973; Shrauger, 1975; and Swann, 1985).

One of the reasons for the inconsistent findings is the focus that researchers take when testing these theories. For example, Shrauger (1975) pointed out that the empirical evidence that supports self-enhancing theory has generally assessed individuals' affective reactions to success or failure. On the other hand, the research that supports self-consistency theory has relied on cognitive reactions to success or failure. Affective reactions would be characterized as emotional ones, such as feelings or preferences for or against something. Cognitive reactions on the other hand entail more thoughtful or rational responses, such as predictions, recall, and judgments of accuracy (McFarlin and Blascovich, 1981). Shrauger (1975) proposed that people with a low level of self-esteem may be more likely to accept and believe negative feedback since it is consistent with their cognitive structures and expectations, even though they

prefer positive feedback because of its affective quality. Given this, we could predict that in terms of affectivity, individuals low in self-esteem may prefer or wish to receive positive feedback, even more so than high self-esteem individuals (self-enhancement). In terms of cognition however, individuals low in self-esteem may predict that they are most likely to receive negative feedback rather than positive feedback (self-consistency).

In order to test this hypothesis, McFarlin and Blascovich (1981) examined both cognitive and affective reactions within the context of the same study. In their study, female subjects' of high, medium and low self-esteem were asked a) their preference for a positive performance evaluation (affective), b) their perceived performance ability (cognitive), and c) their predictions of actual performance (cognitive) on a future task. Results revealed that subjects preferred future success to future failure regardless of their chronic level of self-esteem. In addition, subjects with high and low chronic levels of self-esteem perceived their ability of future performance and expected actual future performance in a manner consistent with their chronic levels of self-esteem. Therefore, these results seem to suggest that although individuals prefer success to failure affectively (self-enhancement theory), their cognitive reactions continue to anticipate outcomes in a manner that is consistent with their chronic level of self-esteem (consistency theory).

Swann, Griffin, Predmore, and Gaines, (1987) made a second attempt to resolve the inconsistencies in the self-evaluation literature. They presented individuals with



high and low levels of self-esteem with either favorable or unfavorable social feedback. They then measured affective reactions to the feedback by assessing mood. In addition, they measured cognitive reactions to the feedback by assessing the participant's reactions to the accuracy of the feedback, competence of the evaluator, diagnosticity of the evaluation technique, and attributions regarding the cause of the feedback. As predicted by self-consistency theory, participants with negative self-concepts indicated that unfavorable feedback was more self-descriptive than favorable feedback. In addition, those participants with negative self-concepts also felt more depressed, anxious, and hostile after they received negative feedback, just as self-enhancement theory would predict. Therefore, the data supported Shrauger's (1975) hypothesis that cognitive reactions to social feedback conform to self-consistency theory and affective reactions conform to self-enhancement theory. However, how would this cognitive versus affective distinction operate in a MSF process?

### Cognitive

Lets imagine an employee that is about to participate in a multi-source feedback performance appraisal. This focal individual is in a position to choose who amongst his or her many peers and subordinates will rate his or her behavior. Based on the theories and studies reviewed thus far, we would expect that cognitively this person would predict, recall and judge as most accurate those evaluation outcomes that are consistent with their level of OBSE. Consequently, individuals low in OBSE would predict that they are more likely to receive a negative performance evaluation. They would also be

more likely to recall negative evaluations and to judge negative evaluations of themselves as most accurate. Individuals high on OBSE on the other hand would predict that they are most likely to receive a positive evaluation, to recall positive evaluations better and to judge positive evaluations of their behavior as most accurate. Therefore, there would likely be a positive relationship between individual level of OBSE and the extent to which those individuals expect positive evaluations, recall past favorable evaluations and judge positive evaluations of their performance as most accurate.

However, in discussing the selection of raters in a MSF process we are not interested in predictions, recall, or judgments of accuracy. In a MSF process that allows individuals to choose the raters of their performance, ratees have the opportunity to influence the process and act upon preferences in order to receive a positive evaluation. According to Shrauger's (1975) taxonomy this ability to act upon a preference would classify rater selection behavior as being affective.

### Affective

Given that affectivity is most salient when choosing raters, the self-enhancement motive will be most active. Individuals will attempt to enhance their self-views especially if they possess a low level of OBSE. Individuals low in OBSE will expect a negative evaluation, and as such will be more likely to manipulate the situation to his or her advantage and choose positive raters (compensatory self-enhancement). Although

individuals high in self-esteem are likely to choose positive raters as well (simple self-enhancement), according to the theories reviewed above, low-self-esteem individuals will be even more likely to do so.

The rationale for this can be further supported by studies which show that individuals low in self-esteem are more likely to have a greater need for approval by others (Kimble & Helmreich, 1972). Given that they negatively evaluate themselves, they may be especially dependent upon others to provide them with positive evaluations. Also, persons with a low level of self-esteem are more likely to view negative feedback as self-diagnostic. That is, they are more likely to over-generalize such feedback to other domains of their identity (Brockner, 1988). For example, after receiving negative feedback about a certain aspect of their work, individuals with low levels of self-esteem are likely to say “I am horrible at my job” or “I am such a complete failure in life”. This psychological process has been labeled “overgeneralization following failure” by Carver and Ganellen (1983). Therefore, negative feedback may have a more severe effect on subsequent efficacy expectations for low self-esteem individuals. These individuals are more likely to feel threatened by the negative information (Kolditz & Arkin, 1982).

In addition, Ashford and Cummings (1983) outlined three motives in the feedback-seeking process that are relevant here: the desire for feedback, the desire to protect one’s ego and the desire to manage one’s impressions. The first of these motives is rational and suggests that individuals will seek useful information to acquire new skills, evaluate abilities, and correct errors in goal-directed behavior. The second

motive results from the fact that feedback is evaluative information and the possibility of receiving negative information about oneself may be threatening to the self-esteem (Ashford, 1986). Lastly, the third motive involves impression management and suggests that feedback seeking is not entirely rational. As was previously discussed, this last motive is especially relevant in the current thesis since it suggests that individuals may engage in activities that influence the nature of the feedback that others give them (Larson, 1989). Employees that engage in impression management may be caught between the need to obtain useful information about their performance and the need to receive a favorable rating in order to present a positive image. Levy, Albright, Cawley, and Williams (1995) used these three motives to build a model of feedback seeking which suggests that a poorly performing employee may be motivated by both ego-defensive concerns and self-presentational concerns to avoid feedback. Avoiding feedback allows them to avoid ego-damaging information and the public embarrassment of being associated with a potentially negative evaluation.

Given all this, it is reasonable to assume that in a MSF process, individuals who have a low level of OBSE will be most likely to choose raters who will rate them positively. OBSE will influence rater selection in that individuals low in OBSE will be especially likely to choose raters who they believe will provide positive evaluations of their behavior at work. However, how will ratees assess who among their many peers and subordinates is most likely to rate them positively? How can “positiveness” be measured? Intuitively, one would think that choosing peers and subordinates that you

are friends with will most likely yield positive feedback evaluations. The next section will review past research that seems to suggest the existence of this relationship.

## **Friendship**

Two key variables that may play a role in determining which raters are likely to give enhanced or positive ratings are likeableness and friendship. These two constructs are very much related and therefore will not be tested separately in this study. Liking can be defined as an affective self-referent evaluative response to a stimulus “represented by a prototype ‘I like Joe’ ” (Zajonc, 1980, p.154). Liking occurs relatively early in the interaction with a stimulus and it is effortless, inescapable and capable of affecting ensuing cognitive processes to a significant degree (Kunst-Wilson & Zajonc, 1980; Moreland & Zajonc, 1979).

Intuitively it makes sense to assume that ratees believe that the raters they like, and are friends with, reciprocate this liking, and are therefore more likely to rate their behavior and attributes in a positive manner. In addition, a substantial amount of evidence supports the notion that liking positively influences outcomes. For example, Cardy & Dobbins (1986) examined if liking operates as an integral dimension in performance appraisals. That is, if it is difficult to separate from other dimensions of performance. In this study, student raters evaluated vignettes that described professors as either likable, dislikable, or neutral in trait terms to determine what effect this had on differential accuracy of performance appraisal evaluations. The results revealed that

liking was indeed an integral dimension of the performance ratings, and that it interfered with a raters' evaluations of performance by increasing noise and error.

These results were replicated by a study conducted by Alexander and Wilkins (1982). In their study, the authors examined the relationship between performance ratings and objective measures of job performance. They found that performance ratings represented the degree of supervisory-subordinate liking (as measured by the amounts of positive social communication between the dyad) to a greater degree than it represented actual quality or quantity of performance (Alexander and Wilkins, 1982). In addition, Dobbins and Russell's (1986) field study found that when subordinates are liked, the leader is less likely to attribute poor performance to ability. This illustrates that performance appraisals are highly influenced by the quality of the relationship between the rater and the ratee.

Additionally, an accumulation of evidence seems to suggest that cold cognitive models of appraisal are inadequate, and that some element of affectivity must be included (Dobbins & Russell, 1986). For example, Dipboye (1985) proposed a model of appraisal that explicitly includes liking as a factor and proposes that it has an independent and direct effect on performance ratings. Tsui and Barry (1986) also assessed the effect that a raters' affect (admiration, respect and liking) towards a ratee has on rating errors (leniency, halo, range restriction, and inter-rater agreement). These authors assessed performance, overall effectiveness and affectivity of individuals in a multidivisional corporation. They found that the ratings provided by raters with positive

affect towards the ratee were the most lenient and that the ratings provided by raters with negative affect were the least lenient. Therefore, the role that liking between ratees and raters may play in influencing appraisals should not be undermined.

In addition, one common reason why organizations may avoid peer assessments is because it is simply perceived as being a popularity contest in which the one with the most friends comes out on top (Love, 1981). It is common sense to assume that those individuals that you like are also your friends and are most likely to provide you with a positive evaluation of your work behavior. Whether this inflation in ratings is intentional or unintentional may vary. However, there are several legitimate reasons to assume it exists. For example, a primary motive of the rater may be to preserve his or her friendship with the ratee, repay personal favors, or avoid a negative situation (Kingstrom and Mainstone, 1985). The effect of all of these motives is an enhanced rating. Since this potential influence may be interpreted by the ratee, it will likely influence who they decide to distribute their evaluation forms to. From this, we can conclude that ratees will perceive friends and individuals they like as the persons who are most likely to give them an enhanced evaluation of their behavior at work. Therefore,

**Hypothesis 1:** Individuals low in OBSE will be more likely than individuals high in OBSE to choose raters who are high in friendship to rate their behavior at work

## **Task Acquaintanceship**

In terms of individuals with a high level of OBSE, we have seen thus far that they view the potential for negative evaluations as less threatening to their egos and are more likely to use that information in a constructive manner to improve performance (McFarlin & Blascovitch, 1981). As such, one could imagine that individuals high in OBSE would be most likely to operate under the self-assessment motive. That is, they would be more likely to seek out an accurate assessment of their work performance. Individuals with a high level of OBSE may be more likely to choose raters who have a high degree of knowledge or acquaintance with their work behavior. The assumption here is that continued interaction and exposure to the target individual's work behaviors should increase the accuracy of the rater's performance evaluation (Farr & Newman, 2000).

For example, Funder and Colvin (1988) found that acquaintanceship led to a higher degree of agreement in personality judgments. Likewise, Moser, Schuler and Funke (1999) found that the validity of assessment center predictors improved considerably once the supervisor and the ratee worked together for more than 2 years. Greller and Herold (1975) found that the "informativeness" of a rating source increased when the source was closer to the individual. Rating sources that work with the ratee often and are acquainted with his/her tasks provide more accurate feedback than a more distant source would. The less contact the rater has with the ratee, the more likely the



responses will be based on rumor and impression and therefore be inaccurate (Maurer & Taurulli, 1996).

Nonetheless, studies supporting the notion that increased task knowledge leads to high validity have been somewhat mixed. For example, some studies have found that increased task acquaintance leads to increased validity (Kornhauser, 1926), increased reliability and less halo error (Ferguson, 1949). Other studies (Jacobs & Kozlowski 1985; Knight, 1923) have found that increased familiarity leads to a larger amount of halo and leniency errors. These inconsistent findings can be somewhat attributed to the different ways in which task acquaintance was measured. The studies that supported the increased task acquaintances, greater accuracy hypothesis made sure to assess task acquaintance separate from personal acquaintance or liking. This issue is especially relevant in this paper since both variables are being explored. As a result, the measure of task acquaintance will be explicitly separated from the friendship measure in this study.

Task acquaintance was assessed along three dimensions: 1) the amount of time the ratee has worked with the peer or subordinate, 2) the degree of contact the ratee has with the target peer or subordinate and 3) the degree to which the peer or subordinate is familiar with the ratee's work processes. These items keep the measure focused on the task and not on the affective reactions that the peer or subordinate has towards the ratee. The use of all three of these items to assess task acquaintanceship is especially important since both the length of time that the rater and ratee have worked together and the

intensity of that interaction are important. For example, Dalessio (1998) suggests a minimum of one year of working together when selecting raters for a given ratee. This is a reasonable benchmark, but the nature of the interaction is also important. Working closely together with a peer for six months may be equivalent to working with a peer for one year in a less intense manner (Farr & Newman, 2000). In addition, there may exist an asymptote past which diminishing returns will be realized for additional opportunities to observe (Farr & Newman, 2000).

Therefore, we can hypothesize that individuals high on OBSE are likely to operate under the self-assessment motive in a multi-source feedback process and choose raters that they feel will rate them accurately. As such,

**Hypothesis 2:** Individuals high in OBSE will be more likely than individuals low in OBSE to choose raters who have a high degree of task acquaintance with their work.

### **Administrative versus Developmental**

In practice, the results of any multi-source feedback process are used for developmental and growth purposes or for administrative and personnel purposes. When feedback is collected for individual development, the feedback is private and is used to highlight strengths and weaknesses with the goal being to improve performance. Under such conditions, the ratee is the only one who receives a copy of the evaluation

and it is kept private. Therefore the ratee owns the data and has full control over who sees the feedback.

When feedback is collected for administrative purposes however, the data is much more public. The immediate supervisor as well as other key constituents within the organization may see the results of the evaluation since it is often used to make decisions concerning compensation, promotion, terminations or lateral transfers (Landy & Farr, 1983). Given this, the use of the evaluation is a contextual variable that can effect employee attitudes towards the process, and thereby the nature of the results.

Morrison and Bies (1991) suggest that the public versus private context of feedback seeking is important because individuals are more likely to engage in defensive impression management and “face-saving” in public settings. Defensive impression management is defined as those behaviors that are intended to avoid creating an unfavorable public self-image. Therefore, in public settings, it is likely that individuals may attempt to manipulate the feedback-seeking environment to ensure that they receive favorable evaluations. In terms of the frequency in which feedback is sought, Ashford and Northcraft (1992) had subjects participate in a computerized task in which they assumed a managerial role and decided whether or not to seek feedback depending on how public the data was. They found that subjects sought feedback less frequently in a public versus a private context. They attributed this to the fact that seeking feedback in public often has increased impression management costs since it may be considered a sign of incompetence or weakness. Ashford and Northcraft (1992)

also found that individuals were more nervous when seeking feedback in public than in private. In addition, Levy, Albright, Cawley and Williams (1995) found that subjects in a private condition sought the most feedback, followed by those in a semi-private condition, and finally those in a public condition. The authors also demonstrated that although subjects may initially decide to seek feedback, they may change their minds if the context also changes from private to public.

However, in addition to the impression management motives that a public context may induce, there are also real losses in pay or status that may result from a negative performance evaluation when the purpose of the appraisal is administrative. This tangible loss, in and of itself may motivate ratees to select raters that will evaluate them positively. For example, Gupta and Jenkins (1996) in their article on politics and pay discuss how some employees may behave differently when performance evaluators are present in an attempt to ensure that they receive a favorable evaluation and all of the benefits that follow. In addition, Kipnis and Venderveer (1971) have demonstrated that employees often use influence tactics, such as referring to their boss as a “nice guy” in order to secure a pay raise. Given this, it is not difficult to imagine that when a MSF process is to be used for administrative purposes and ratees are allowed to choose their raters, they will likely choose those individuals whom they feel will rate them positively. This relationship seems especially likely to occur since selectively choosing potential raters is quite subtle and may not directly point towards ulterior motives (Martin, 1987).

In addition, London, Wohlers and Gallagher (1990) found that 34% of the subordinates in their study rated their bosses differently (more lenient) when the feedback was intended for administrative purposes such as affecting the manager's pay. Therefore, we have seen that the purpose of the appraisal, developmental versus administrative, does influence whether rater biases occur. This suggests that it may just as well influence the occurrence of ratee biases. Therefore, the feedback environment (private versus public) will moderate the self-assessment and self-enhancement motive that is expected to operate with individuals high and low in self-esteem, respectively. When the feedback data is intended to be public and influence pay and promotional decisions, all employees, regardless of their level of self-esteem will attempt to seek out raters who they believe will provide an enhanced evaluation.

This can be supported by the fact that Farr and Newman (2000) suggest that when the evaluation is to be used for administrative purposes, the selection of the specified raters from the eligible list should be under organizational control. The underlying assumption is that the potential for the ratee to manipulate the situation in order to guarantee enhanced ratings is great. Therefore,

**Hypothesis 3:** Regardless of an individual's level of OBSE, when feedback will be used for administrative purposes, individuals will be more likely to seek raters who are high in friendship than when feedback will be used for developmental purposes.

## **Subordinates/Peers**

Given the different work and power dynamics that exist between peers and subordinates, it is possible that ratees may use different strategies when choosing amongst these different rating sources. For example, French and Raven (1959) state that there are five bases of power amongst organizational members: legitimate power, referent power, reward power, expert power, and punishment power. Regardless of the exact amounts or types of power that peers versus subordinates may hold in reference to the ratee, these dynamics do exist. Therefore, different motives may operate when ratees are choosing which peers they would like to rate their performance versus which subordinates they would like to rate their performance.

Maurer, Raju, and Collins (1998) touched upon this issue when they examined the measurement equivalence of peer and subordinate ratings on a team-building skill dimension. The peers and subordinates both rated their managers on this dimension and the results revealed that the two rating groups did not differ significantly. Nonetheless, they did find that peer raters, as compared to subordinate raters, had slightly higher standards. Therefore, it took more skill for a manager to receive a peer rating in the highest category. This suggests that perhaps peers have greater knowledge and experience and therefore possess higher expectations (Maurer, Raju, and Collins, 1998). As a result, they are tougher raters than the subordinates. However, Bettenhausen & Fedor (1997) state that peers have firsthand experience with the external constraints of

the job, therefore peers will be more empathetic in evaluating performance than other sources will. Therefore, it is unclear how ratees may perceive peer evaluations.

Nonetheless, only one researcher has directly examined the patterns of feedback seeking across different sources, or the preference for one source over another (Ashford & Tsui, 1991). Ashford and Tsui (1991) hypothesized that employees may not treat all sources of feedback equally. These authors stated that one factor which has the potential for influencing the use of different strategies is the power-dependence relationship between the employee seeking the feedback and the source (Eder & Fedor, 1989). Therefore, although all managers may depend on their peers and subordinates to varying degrees, their superiors are the ones who control their rewards and therefore their dependence on this group is highest. Given this high level of dependence, Ashford and Tsui (1991) hypothesized that managers would be more active in seeking feedback from superiors than from peers and subordinates. This hypothesis was supported by the data and suggests that employees are especially interested in understanding how their supervisors view them so they can alter their behavior if necessary. If superiors are the ones that control their future rewards and sanctions then it follows that they would like to obtain an accurate evaluation from this group.

In addition, Ashford and Tsui (1991) found that managers would seek more negative (and thereby more accurate) feedback from superiors and subordinates and the least amount of negative feedback from peers. Managers would seek the most positive feedback from superiors and peers and the least positive feedback from subordinates.

This finding is most instrumental for the current research project because it suggests that seeking negative information from peers may incur larger impression management costs and be the most ego threatening when the performance evaluation is public. Receiving negative performance information from superiors may be in line with their functional responsibility to employees (Bettenhausen & Fedor, 1997). Employee may feel that since their superior is more experienced and knowledgeable about the job, negative feedback is appropriate from this source and can lead to employee development by correcting performance. On the other hand, negative performance information from peers may be regarded as more ego threatening. Peers are often closer in age to the focal employee, they serve as a social reference group and they often compete for rewards. Therefore, receiving critical or negative evaluations from this in-group may be shunned (Bettenhausen & Fedor, 1997). Given this, employees may actively avoid potentially negative information from their peers.

As for subordinates, Ashford and Tsui (1991) suggested that the employees may have been more willing to seek negative evaluations from this group in order to portray themselves as interested, responsive and caring managers. In addition, receiving a negative evaluation from subordinates may be less ego threatening since they do not serve as a reference group, nor do they compete for rewards.

While this study is useful in suggesting the influence that the rating source may have on the ratee's selection of raters in a MSF process, the study does not discuss self-esteem. In the current study, the impact that the rating source (peer versus subordinate)



may have in the selection of raters will be examined but it will be exploratory, and no hypotheses will be stated.

## **METHOD**

### **Design**

The design of this study is a 2 (source = peer versus subordinate) X 2 (purpose = administrative versus developmental) X 2 (selection = selected versus not selected) within subject, fully crossed repeated measures design with one continuous independent variable (OBSE). The design is fully crossed, which means that every participant was exposed to every level of each independent variable. A principle advantage of repeated measures designs is that they provide good precision for comparing treatments because all sources of variability between subjects are excluded from the experimental error. Therefore, the participants in the study serve as their own controls. The dependent variables are: 1) level of friendship and 2) task acquaintanceship. The data collection method used was a web-based survey.

### **Sample**

The sample consisted of 91 MBA and Executive MBA students from 8 Universities across Canada. The director of the MBA and EMBA program in each major university across Canada was contacted and his or her help in this research project was solicited (see Appendix 1 for the message). Therefore, sixteen MBA directors were contacted and 8 decided to participate, yielding a response rate of 50%. The MBA directors were asked to forward an e-mail message to all of their MBA and EMBA students. The message solicited completion of the survey. It detailed the persons

responsible for the research, the time required to complete the survey, the fact that the survey data would remain anonymous and the web-page address. In addition, the message asked that the survey only be completed by individuals who currently worked full-time or had worked full-time in the past two years and had a minimum of four subordinates under their supervision (see Appendix 2 for the message). In total, approximately 1,500 MBA and Executive MBA students were sent an e-mail. It is difficult to know exactly how many students received and read the e-mail however. Out of these 1,500 students, 91 fully completed the survey and their data was used in this study. This yields an approximate response rate of .06%. In addition, 32 data entries were eliminated from the analyses because of incomplete responding.

Data collection occurred in two waves. Within each wave of data collection an initial e-mail was sent to the students asking them to participate in the survey and then two reminder e-mails were sent at one-week intervals in order to help boost the response rate (Summers & Groehler, 2000). The first wave occurred in the summer month of July. The first wave of data collection yielded a very low response rate of approximately .02%. This low response rate may have been due to the following factors. First, because the e-mail was sent in the summer month of July students might have been out of town and not checking their school e-mail accounts. Second, the survey was long and no incentive to complete the survey was offered. Therefore, the second wave of data collection occurred in the school month of September and a draw for a \$150 gift certificate was included. This second wave of data collection did yield a better response rate of .05%. Nonetheless, this response rate was still very low.

The mean age of the respondents in this study was 34 years ( $SD=6.38$ ). The mean number of years the respondents spent in school (from elementary onwards) was 20 ( $SD=3.00$ ). Sixty-two percent of the respondents were male and 38% were female. Half of the respondents held a position of director or manager (49%). While the other respondents occupied the positions of engineer (6.5%), coordinator (5.5%), vice-president (4%), supervisor (4%), owner (3.5%), controller (3.5%), team leader (3.5%), financial advisor (2%), and “other” (18.5%). In terms of organizational level, 11% of the respondents occupied top management positions, 8.8% were executives, 18.7% were upper management, 46.2% were middle management, 13.2% were first level management and 2.1% classified themselves as “other”. With regards to work function, 14.3% were in operations, 14.3% were in credit/finance, 11% were in engineering, 8.8% were in HR/training, 8.8% were in administration, and 42.8% occupied a variety of other work functions. In terms of the highest degree earned, 53.8% earned a bachelor’s degree, 45.1% a master’s degree, and 1.1% a doctorate/professional degree. Twenty-three percent of respondents worked in transportation, communication and utilities, 23% worked in finance, insurance and banking, 11% worked in manufacturing, while the other 43% worked in a variety of other organizations.

### **Experimental Manipulation**

The survey first assessed the respondent’s level of OBSE. Next, it asked respondents to name four peers and four subordinates that they worked with. Respondents then answered a series of descriptive questions pertaining to each peer and

subordinate they named. These questions assessed the perceived level of friendship and task acquaintanceship. Lastly, the respondent was asked to indicate which two peers and which two subordinates they would choose to rate their behavior if the performance evaluation was to be used for administrative purposes and which two they would choose if the evaluation was to be used for developmental purposes.

Although Chappelow (1998) stated that in most MSF processes the typical number of raters required is five peers and five subordinates, it was not realistic to have the respondents of our survey describe this many raters. In order to fit with the typical practice, respondents would have had to describe approximately ten potential raters and then choose five out of the ten described. This would have made the survey too long and we therefore opted for the description of four and the choice of two.

### **Pilot Test**

Given the novelty of this research method (i.e., use of a web-based survey) and the fact that the items of some of the measures were modified in some manner, the survey was pilot tested. Seventy-five contacts were e-mailed a message requesting completion of the web-based survey. Eighteen individuals completed the survey and provided comments about both the content and the layout of the survey. The comments helped make the survey more user-friendly and easy to complete. One example of a suggestion that was incorporated into the design of the survey was the addition of a page confirming the completion of the survey once the respondents pressed the submit button.

In addition, a reliability analysis of the measures employed in the survey was conducted in order to ensure that the items had high internal consistency. These results will be presented in the measures section below.

### **Ethical Provisions**

The project was reviewed and approved by the Concordia Ethics committee for research involving human subjects. All data remained anonymous in that there was no way of identifying participants in the study. All participants were allowed to terminate participation in the study at any time by simply exiting the web-page.

### **Web-Based Survey**

In order to test the hypotheses stated above a web-based survey was developed (see appendix 3-Part 1-6). The primary reason behind the selection of this method was because it allowed the survey to be interactive. The web-based survey was designed so that information entered at the beginning of the survey was able to reappear in questions later in the survey. For example, respondents in our survey were asked to describe four of their peers. Since they provided the names of their peers early in the survey, the descriptive questions were later tailored to each specific peer. For example, “how frequently do you have contact with *Susan*”? This was especially important in our survey since there were eight descriptive questions for each of the four peers and four

subordinates that they chose to describe. Inserting the name of the focal individuals into the question allowed the respondent to stay focused on that individual.

In addition, a web-based survey provides cleaner data since prompts can be programmed that eliminate the possibility of out-of-range or unacceptable values. For example, respondents completing our web-based survey were asked to choose two subordinates among four. The program would not allow them to move on if they chose only one, three, or four. A web-based survey also points the respondent toward an unanswered question before moving on, so there is less missing data.

As with any data collection method, some issues are inherent in the use of the Internet for data collection, these are 1) sampling problems, 2) response consistency problems and 3) participant motivation problems (Stanton, 1998). The first problem refers to the fact that all documents published on the Internet are freely available to anyone who happens to stumble upon them. Therefore, anyone other than the individuals contacted to participate in this study may have completed the survey and submitted data. Although the potential for this is real, the likelihood of it occurring in an undetected manner was very small. The questions that asked about the years of schooling completed and the highest degree earned in the demographic section at the beginning of the survey served as a check that only the MBA students contacted completed the survey. Given that the target sample consists of MBA students, a minimum of approximately 16 years of completed schooling must have been entered (from elementary school onwards) and all respondents must have indicated that the

highest degree earned to date was a bachelor's degree. In addition, incomplete data sets and erroneous or fixed responding to questions as well as erroneous names of peers and subordinates were also indicators that the survey was not being completed seriously.

Second, the respondent has full control over the time and setting in which the survey is completed. Therefore, the respondent's psychological state may vary more than in a controlled experimental setting. Nonetheless, it would not be any different from a paper and pencil based survey.

Third, respondents may lose interest in the survey while they are completing it, since there is no experimenter present to encourage completion. This problem is one that can occur in any study that does not enforce participation.

Nonetheless, a study conducted by Stanton (1998) used identical questionnaire items to gather data from two samples of employees. One sample responded to the questionnaire via the Internet and the second sample responded to the questionnaire via the traditional paper and pencil format. The results revealed that the amount of missing data was lower for data records collected using the Internet. In addition, the factor structure of the items forming a scale and the correlations between the scales was not different across the data collection formats (Stanton, 1998). Nonetheless, because the paper and pencil and web respondents were employees from different types of organizations and the sample size was relatively small, this study was replicated by Young, Daum, Robie and Macey (2000). These researchers used over 1,500



respondents from the same organization to examine the differences between web-based and paper and pencil surveys. Again, they found that the amount of missing data was lower for the web sample. In addition, the descriptive statistics revealed virtually no differences in the results of the two samples and differential item and test functioning did not result in any consistent relationship between mode of response and response patterns. Therefore, both studies found similar results and both suggest that web-based surveys yield similar results to paper and pencil surveys. This provides encouraging evidence for the quality and usefulness of data collected over the Internet.

## **Measures**

### **Demographic variables**

The demographic variables assessed were: age, number of years in school, gender, position held, organizational level, organizational function, highest degree earned, and type of organization the respondents work for (see appendix 3-Part 1).

### **Organizational based self-esteem**

The measure of OBSE developed by Pierce, Gardner, Cummings, and Dunham (1989) was used. Each item reflects the extent to which employees believe they are valuable, worthwhile and effectual members of their employing organizations. The respondents were asked to think about the messages they received from the attitudes and

behaviors of their managers and supervisors and to indicate on a 5-point scale the extent to which they agreed or disagreed with each of the ten statements. The items were slightly modified in this study so they were more context-specific. Therefore, rather than stating I count *around here* or I am taken seriously *around here*, the items were tailored to be specific to their current or previous job. For example, I count *at XYZ* or I am taken seriously *at XYZ* (see appendix 3-Part 2).

In Pierce et al's (1989) sample, the alpha coefficient of the OBSE measure was equal to 0.86 or greater, ranging to a high of 0.96 in one sample. The pilot study conducted in this study yielded an alpha coefficient of 0.93, and the data collected for the main analyses yielded an alpha value of 0.91. The strength of these internal consistency estimates provides evidence for the homogeneity of the scale items.

#### Assessments for each peer and subordinate

The next part of the web-based survey asked respondents to name four subordinates and four peers that they worked with. Once they did that, they were asked a series of eight questions for each peer and subordinate. The questions are detailed below in the friendship and task acquaintanceship section.

## Friendship

The four items developed by Wayne and Ferris (1990) to measure likeability were used as a template for measuring friendship and likeability in this study. The three likeability items remained the same; the only difference being whether they targeted a peer or a subordinate (see appendix 3-Part 3 & 4). The one friendship item was modified from “I think this peer or subordinate would make a good friend” to: “I can consider this peer or subordinate a good friend” in order to get a more direct measurement of friendship. The items were assessed on a 5-point Likert scale and these questions were repeated for every focal peer and subordinate. In addition, a question pertaining to the perceived valence of the rating was assessed by directly asking: “If this person were to rate your performance, what kind of evaluation would they give you overall”. The anchors ranged from (1) highly positive to (5) highly negative (see appendix 3-Part 3 & 4).

The pilot data revealed that the four likeable and friendship items had an alpha coefficient of 0.91 and that value increased to 0.93 once the direct valence item was added in. The fact that the inclusion of the valence item increased the alpha supports the contention that friends and liked peers or subordinates are perceived to be more likely to rate the focal individual positively during a performance evaluation. In the main analyses, the friendship items yielded an alpha of 0.92, demonstrating that they are indeed assessing one construct.

### Task acquaintanceship

Task acquaintanceship was assessed with 3 items, each taken from different sources. Item one asked: “how long have you worked with or supervised this peer or subordinate” and was taken by Knight (1923). The second item asked: “how frequently do you have contact with this individual”; this item was taken from Freeberg (1969) and the anchors ranged from (1) very frequently to (5) not frequently at all. The third item asked: “how familiar is he or she with your work performance”. This item was taken from Ferguson (1949) and the anchors ranged from (1) very familiar to (5) very unfamiliar. The three items were used here in order to avoid any problems associated with one-item measures (see appendix 3-Part 3 & 4).

The reliability analysis assessing the internal consistency in these items from the pilot test was quite low, yielding an alpha coefficient of 0.47. This value increased to 0.77 however when the “how long have you worked with this peer or subordinate” item was removed. A similar pattern of results was found for the reliability analysis of the main data collected in this study. The 3 task acquaintanceship items yielded an alpha of 0.42, but this alpha increased to 0.76 when the “length of time” item was removed. The assumption behind the use of the “length of time” item is that working with someone for a longer period of time should lead to higher task acquaintanceship. However, this is not necessarily true since working with someone closely for one month, every day, can lead to higher task acquaintanceship than working with someone for one year on an occasional basis. Therefore, this item may indeed be assessing something separate from

task acquaintanceship. Consequently, this item was analyzed separately from the other task items.

### Selection of raters

In order to determine which peers or subordinates a respondent would choose to rate their behavior, the last part of the questionnaire asked the following:

“Now, imagine that you had to choose 2 out of the 4 peers or subordinates you described above to evaluate your performance. Which 2 would you choose if the ratings were going to be private, that is, no one but yourself (not even your immediate supervisor) would see the results? Which 2 peers or subordinates would you choose if the ratings were to be public, that is, the results would be available to your supervisor and could influence decisions related to compensation, promotion, terminations or lateral transfers”?

Below each question, the four names of the peers or subordinates that they described in the survey were listed and they had to choose two out of the four. (See appendix 3-Part 5 & 6).

## ANALYSIS

### Descriptive Analyses

A summary of the variables in this study is presented in Table 1 below.

**Table 1: Means, standard deviations and correlations (N = 91)**

	Mean	SD	1	2	3	4	5	6	7	8
1)OBSE	4.23	0.60	1.00							
<b><u>FRIEND</u></b>										
2)Adm.	3.76	0.52	.22*	1.00**						
3)Dev.	3.76	0.52	.22*	1.00**	1.00					
4)Sel.	3.94	0.49	.21*	.94**	.94**	1.00				
5)Not Sel.	3.59	0.61	.20	.96**	.96**	.82**	1.00			
6)Peer	3.81	0.56	.06	.78**	.78**	.71**	.76**	1.00		
7)Sub.	3.71	0.71	.27**	.87**	.87**	.83**	.83**	.37**	1.00	
<b><u>TASK</u></b>										
8)Adm.	3.81	0.43	.34**	.66**	.66**	.66**	.60**	.51**	.58**	1.00
9)Dev.	3.81	0.43	.34**	.66**	.66**	.66**	.60**	.51**	.58**	1.00**
10)Sel.	3.88	0.49	.28**	.55**	.55**	.56**	.50**	.42**	.48**	.93**
11)Not Sel.	3.74	0.45	.34**	.68**	.68**	.67**	.62**	.52**	.59**	.92**
12)Peer	3.87	0.50	.24*	.51**	.51**	.49**	.48**	.48**	.37**	.87**
13)Sub.	3.74	0.50	.33**	.64**	.64**	.66**	.56**	.39**	.63**	.87**
Length of Time										
14)Adm.	2.87	1.98	.21*	.24*	.24*	.21	.25*	.11	.27*	.35**
15)Dev.	2.87	1.98	.21*	.24*	.24*	.21	.25*	.11	.27*	.35**
16)Sel.	2.84	1.89	.23*	.20	.20	.16	.21*	.07	.24*	.35**
17)Not Sel.	2.89	2.18	.19	.26*	.26*	.23*	.26*	.14	.27**	.32**
18)Peer	3.15	2.34	.23*	.21*	.21*	.17	.23*	.11	.23*	.33**
19)Sub.	2.58	1.88	.16	.24*	.24*	.22*	.24*	.11	.27**	.32**

**Table 1 (Continued): Means, standard deviations and correlations (N = 91)**

	9	10	11	12	13	14	15	16	17	18	19
1)OBSE											
<b><u>FRIEND</u></b>											
2)Adm.											
3)Dev.											
4)Sel.											
5)Not Sel.											
6)Peer											
7)Sub.											
<b><u>TASK</u></b>											
8)Adm.											
9)Dev.	1.00										
10)Sel.	.93**	1.00									
11)Not Sel.	.92**	.71**	1.00								
12)Peer	.87**	.81**	.79**	1.00							
13)Sub.	.87**	.80**	.80**	.50**	1.00						
Length of Time											
14)Adm.	.35**	.37**	.26*	.28**	.32**	1.00					
15)Dev.	.35**	.37**	.26*	.28**	.32**	1.00**	1.00				
16)Sel.	.35**	.38**	.27*	.29**	.32**	.97**	.97**	1.00			
17)Not Sel.	.32**	.35**	.24*	.26*	.30**	.98**	.98**	.89**	1.00		
18)Peer	.33**	.37**	.23*	.27**	.29**	.95**	.95**	.92**	.93**	1.00	
19)Sub.	.32**	.33**	.27*	.25*	.31**	.92**	.92**	.89**	.90**	.75**	1.00

NOTE : Variable names : OBSE = Organizational based self-esteem, Adm = Administrative/Public  
Dev = Development/Private, Sel = Selected, Not Sel = Not Selected, Peer = Peer, Sub = Subordinate

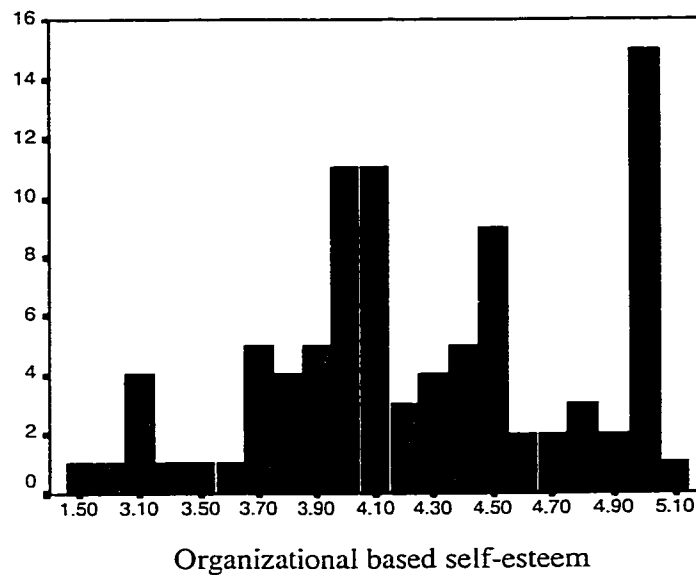
\* Correlation is significant at the 0.05 level (2-tailed)

\*\* Correlation is significant at the 0.01 level (2-tailed)

The means, standard deviations and correlations amongst the variables are reported in the table. The variables are organized along the two main dependent variables (friendship and task acquaintanceship). Beneath the task acquaintanceship variable is a separate variable labeled length of time. This variable represents one item in the survey that asked respondents to detail how long they worked with the particular peer or subordinate they were describing. As stated earlier, although this item was intended to be a task acquaintanceship item, inter-item reliability analyses demonstrated that it did not fit well with the other task items. Therefore, this variable was analyzed separately from the other task items.

From the table we can see that the mean OBSE score is high (4.23 out of a possible 5). In addition, the standard deviation of the OBSE variable is not large (SD=0.60), indicating that there is little variance in the OBSE scores. A plot of the OBSE scores demonstrates this (see Figure 1 below).

**Figure 1: OBSE Frequency**



A measure of skewness was performed on the OBSE variable and the skewness statistic equaled  $-1.151$ . This skewness statistic indicates that the OBSE variable is negatively skewed and not normally distributed. This violates the assumption of normality for multivariate statistics (Tabachnick & Fidell, 1996).

Two data transformations were conducted on the OBSE variable (Log10 and Square Root) to reduce the negative skew. However, these transformations had little



effect, a somewhat common occurrence (Wike & Church, 1982). Therefore, the OBSE variable will be analyzed as is.

### **Main Analyses: Tests of the hypotheses**

The design of this study is a fully crossed within subjects repeated measures with one continuous independent variable (OBSE). The raw data collected from the survey was aggregated by calculating the mean friendship and the mean task acquaintanceship score for each combination of the 3 conditions (source, purpose and selection). The source condition has 2 levels (peer and subordinate), the purpose condition has 2 levels (administrative and development) and the selection condition also has 2 levels (selected and not selected). Given that each condition had 2 levels, 16 values were calculated (8 for friendship and 8 for task acquaintanceship), as demonstrated in Tables 2 & 3 below:

**Table 2: Friendship Variables**  
**FRIENDSHIP VARIABLES**

	Administrative		Development	
	Selected	Not Selected	Selected	Not Selected
Peer	FASP	FANP	FDSP	FDNP
Subordinate	FASS	FANS	FDSS	FDNS

**Table 3: Task Acquaintanceship Variables**

TASK ACQUAINTANCESHIP VARIABLES				
	Administrative		Development	
	Selected	Not Selected	Selected	Not Selected
Peer	TASP	TANP	TDSP	TDNP
Subordinate	TASS	TANS	TDSS	TDNS

For example, the variable FASP in Table 2 is the mean friendship score(F) when the purpose of the appraisal is administrative(A) for selected(S) peers(P). Likewise, the variable FANP in Table 2 is the mean friendship score(F) when the purpose of the appraisal is administrative(A) for not selected(S) peers(P).

Therefore, in the analyses performed in this study, the mean friendship and task acquaintanceship scores for both the selected AND the not selected raters was calculated. It was important to do this so that we could compare those raters who were selected against those who were not. Breaking down the data in this manner meant that the data had to be analyzed separately for each purpose condition (administrative and developmental) since the overall friendship and task acquaintanceship means of each condition would be equal (see Tables 4 & 5 below).

**Table 4: Mean Friendship Score for each condition**

PURPOSE = Administration (mean = 3.761)			
SOURCE			
PEER (Mean = 3.809)		SUBORDINATE (Mean =3.713)	
Peer Selected (Mean = 4.022)	Peer Not Selected (Mean = 3.596)	Subordinate Selected (Mean = 3.879)	Subordinate Not Selected (Mean = 3.546)
PURPOSE = Developmental (mean = 3.761)			
SOURCE			
PEER (Mean =3.809)		SUBORDINATE (Mean =3.713)	
Peer Selected (Mean = 4.019)	Peer Not Selected (Mean = 3.598)	Subordinate Selected (Mean = 3.824)	Subordinate Not Selected (Mean = 3.601)

**Table 5: Mean Task Acquaintanceship Score for each condition**

PURPOSE = Administration (mean =3.806)			
SOURCE			
PEER (Mean =3.870)		SUBORDINATE (Mean =3.742)	
Peer Selected (Mean = 3.868)	Peer Not Selected (Mean = 3.871)	Subordinate Selected (Mean = 3.860)	Subordinate Not Selected (Mean = 3.623)
PURPOSE = Developmental (mean = 3.806)			
SOURCE			
PEER (Mean =3.870)		SUBORDINATE (Mean =3.742)	
Peer Selected (Mean = 3.957)	Peer Not Selected (Mean = 3.782)	Subordinate Selected (Mean = 3.815)	Subordinate Not Selected (Mean = 3.668)

From the table we can see that the overall mean for friendship in the administrative condition (3.761) is equal to the overall mean for friendship in the developmental condition (3.761). Analyzing the two purpose conditions (administrative and developmental) together would have yielded a perfect correlation for the main effect of purpose (since the two means are identical) and therefore, would not allow us to test for any interaction effects involving purpose. As a result, the data was analyzed separately for the 2 levels of purpose. In addition, the data was analyzed separately for the two dependant variables, friendship and task acquaintanceship and the “length of

time” item. Repeated measures ANCOVA analyses were conducted since each respondent in the study was exposed to every level of the independent variable and OBSE acted as the covariate.

### Friendship

From Table 6 below we see that the main effect for selection was not significant [ $F(1, 89) = 2.25, p=ns$ ] for the administrative condition. In addition, the main effect for selection was not significant [ $F(1, 89) = 0.98, p=ns$ ] for the developmental condition (Table 7). This means that regardless if the purpose of the appraisal was administrative or developmental, the selected raters were not significantly different from the not selected raters in terms of perceived friendship. Therefore, hypothesis 3, which stated that when the purpose of the appraisal was administrative, individuals would be more likely to choose raters who they perceived as being higher in friendship, was not supported. Support for this hypothesis would have been evidenced if there was a significant main effect for selection in the administrative condition (Table 6) but not in the developmental condition (Table 7).

**Table 6: Repeated Measures Results for the Administrative Condition:  
Friendship Outcome**

		Overall F	DF	Significance	Eta. Squared
	Selection	2.25	(1,89)	0.137	0.025
	Source	5.14	(1,89)	0.026**	0.055
	Selection * OBSE ( $H_{01}$ )	0.16	(1,89)	0.686	0.002
	Source * OBSE	4.44	(1,89)	0.038**	0.047
	Source * Selection	0.01	(1,89)	0.911	0.000
	Source * Selection * OBSE	0.00	(1,89)	0.991	0.000

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*( $H_{01}$ =hypothesis 1)*

**Table 7: Repeated Measures Results for the Developmental Condition:  
Friendship Outcome**

		Overall F	DF	Significance	Eta. Squared
	Selection	0.98	(1,89)	0.325	0.011
	Source	5.14	(1,89)	0.026**	0.055
	Selection * OBSE ( $H_{01}$ )	0.07	(1,89)	0.787	0.001
	Source * OBSE	4.44	(1,89)	0.038**	0.047
	Source * Selection	0.53	(1,89)	0.469	0.006
	Source * Selection * OBSE	0.19	(1,89)	0.666	0.002

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*( $H_{01}$ =hypothesis 1)*

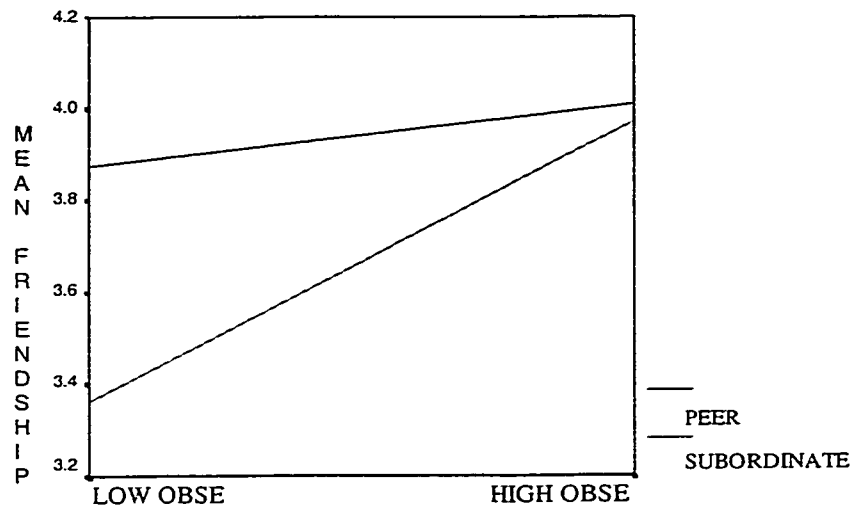
In terms of the main effect for source, the mean friendship score for peers (3.809) is compared to the mean friendship score for subordinates (3.713). These means remain the same for both the administrative and developmental purpose (refer to Table 4 above). This occurs because the main effect of source does not include the selection manipulation and simply compares the mean friendship score for peers (both those selected and those not selected) against the mean friendship score for subordinates (both those selected and those not selected) therefore it cannot differ across purpose. From Tables 6 and 7 we see that the main effect for source was significant [ $F(1, 89) = 5.14, p < 0.05$ ]. This means that peers and subordinates differed significantly in terms of

perceived friendship. The mean friendship score for peers was 3.809 (SD = .56) and the mean friendship score for subordinates was 3.713 (SD = .71). Therefore, respondents perceived their peers as being significantly higher in friendship than subordinates.

The test for hypothesis 1 can be found by examining the selection by OBSE interaction. As demonstrated in Tables 6 and 7 above, the selection by OBSE interaction was not significant when the purpose of the appraisal was administrative [ $F(1, 89) = 0.16, p=ns$ ] and when the purpose of the appraisal was developmental [ $F(1, 89) = 0.07, p=ns$ ]. Therefore, hypothesis 1, which stated that individuals low in OBSE would be more likely than individuals high in OBSE to choose raters who they perceived as being high in friendship to rate their behavior at work, was not supported in either purpose condition.

The interaction effect of source by OBSE however was significant [ $F(1,89)=4.44, p<0.05$ ] (Tables 6 & 7). This means that respondent's level of OBSE played a role in influencing the perceived friendship of peers and subordinates. Again, the means used in the analysis of this interaction and therefore the results are exactly the same for both administrative and developmental purposes since the selection manipulation is not included. The mean friendship score for peers (both those selected and those not selected) and subordinates (both those selected and those not selected) are compared and therefore they cannot differ across purpose (see Table 4). The source by OBSE interaction is illustrated in Figure 2 below.

**Figure 2: Source by OBSE interaction for friendship**



In order to display the source by OBSE interaction, the sample was dichotomized between low and high OBSE. Low OBSE was represented by one standard deviation point below the mean and high OBSE by one standard deviation point above the mean. From the figure we can see that those respondents who were low in OBSE made a greater differentiation in the perceived friendship of their peers versus their subordinates. They perceived their subordinates to be significantly lower in friendship than their peers. In contrast, individuals high in OBSE did not make such a large distinction between the perceived friendship of their peers versus their subordinates.

In terms of the exploratory analyses involving rater source, the question was: do individuals select raters differently depending on whether they are choosing a peer versus a subordinate? The answer to this question can be found by examining the source by selection interaction in Tables 6 and 7. The source by selection interaction

was not significant for both the administrative purpose [ $F(1, 89) = 0.01, p=ns$ ] and the developmental purpose [ $F(1, 89) = 0.53, p=ns$ ]. Respondents therefore did not select peers versus subordinates differently depending on their perceived level of friendship. Although the main effect for source illustrated that peers were perceived as being better friends than subordinates, this did not seem to determine the selection of raters.

The source by selection by OBSE interaction was also not significant for both the administrative purpose [ $F(1,89)=0.00, p=ns$ ] and the developmental purpose [ $F(1,89)=0.19, p=ns$ ] (Tables 6 & 7). Therefore, respondent's level of OBSE did not influence the selection of peers versus subordinates depending on their level of perceived friendship. Overall these results suggest that individuals do not choose peers versus subordinates differently based on the level of friendship they have with them. In addition, OBSE plays no role in influencing this relationship.

#### Task Acquaintanceship

From Table 8 below we see that the main effect for selection was not significant [ $F(1, 89) = 0.02, p=ns$ ] for the administrative condition. In addition, the main effect for selection was not significant [ $F(1, 89) = 1.34, p=ns$ ] for the developmental condition (Table 9). This means that regardless if the purpose of the appraisal was administrative or developmental, the selected raters were not significantly different from the not selected raters in terms of perceived task acquaintanceship.



**Table 8: Repeated Measures Results for the Administrative Condition:  
Task Acquaintanceship Outcome**

		Overall F	DF	Significance	Eta. Squared
	Selection	0.02	(1,89)	0.884	0.000
	Source	1.50	(1,89)	0.224	0.017
	Selection * OBSE ( $H_{o2}$ )	0.03	(1,89)	0.861	0.000
	Source * OBSE	0.80	(1,89)	0.375	0.009
	Source * Selection	0.03	(1,89)	0.853	0.000
	Source * Selection * OBSE	0.02	(1,89)	0.900	0.000

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*( $H_{o2}$ =hypothesis 2)*

**Table 9: Repeated Measures Results for the Developmental Condition:  
Task Acquaintanceship Outcome**

		Overall F	DF	Significance	Eta. Squared
	Selection	1.34	(1,89)	0.250	0.015
	Source	1.50	(1,89)	0.224	0.017
	Selection * OBSE ( $H_{o2}$ )	0.61	(1,89)	0.435	0.007
	Source * OBSE	0.80	(1,89)	0.375	0.009
	Source * Selection	0.73	(1,89)	0.396	0.008
	Source * Selection * OBSE	0.68	(1,89)	0.411	0.008

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*( $H_{o2}$ =hypothesis 2)*

From Tables 8 and 9 we see that the main effect for source was not significant [ $F(1, 89) = 1.50, p=ns$ ]. This means that peers and subordinates did not differ significantly in terms of perceived task acquaintanceship. Again, the means used in this analysis and therefore the results are exactly the same for both administrative and developmental purposes since there is no selection manipulation.

The test for hypothesis 2 can be found by examining the selection by OBSE interaction. As demonstrated in Tables 8 and 9 above, the selection by OBSE

interaction was not significant when the purpose of the appraisal was administrative [ $F(1, 89) = 0.03, p=ns$ ] and when the purpose of the appraisal was developmental [ $F(1, 89) = 0.61, p=ns$ ]. Therefore, hypothesis 2, which stated that ratees high in OBSE would be more likely than ratees low in OBSE to choose raters who they perceived as having a high degree of task acquaintanceship with their work, was not supported in either purpose condition.

The interaction effect of source by OBSE was not significant [ $F(1,89)=0.80, p=ns$ ] (Tables 8 & 9). This means that respondent's level of OBSE did not play a role in influencing the perceived task acquaintanceship of peers and subordinates.

In terms of the exploratory analyses involving rater source, the source by selection interaction was not significant for both the administrative purpose [ $F(1, 89) = 0.03, p=ns$ ] and the developmental purpose [ $F(1, 89) = 0.73, p=ns$ ] (Tables 8 & 9). Respondents did not select peers versus subordinates differently depending on their perceived level of task acquaintanceship.

The source by selection by OBSE interaction was also not significant for both the administrative purpose [ $F(1,89)=0.00, p=ns$ ] and the developmental purpose [ $F(1,89)=0.19, p=ns$ ] (Tables 8 & 9). Therefore, respondent's level of OBSE did not influence the selection of peers versus subordinates depending on their level of perceived task acquaintanceship. Overall these results suggest that individuals do not choose peers versus subordinates differently based on how acquainted they may be with

the tasks they perform at work. In addition, OBSE plays no role in influencing this relationship.

### Length of time

As discussed previously, the length of time variable, which assessed how long (in years) the ratee has worked with the potential rater, was analyzed separately from the other task items. As we can see in Tables 10 and 11 below, no main effect for selection was found for the administrative condition [ $F(1,89)=0.76, p=ns$ ] or the developmental condition [ $F(1,89)=0.36, p=ns$ ]. In addition, the selection by OBSE interaction was not significant for the administrative condition [ $F(1,89)=0.88, p=ns$ ] or the developmental condition [ $F(1,89)=0.53, p=ns$ ]. Therefore, hypothesis 2 was not supported. These results are very similar to the results obtained for the task acquaintanceship scale items (Tables 8 & 9).

**Table 10: Length of Time Repeated Measures Results for the  
Administrative Condition**

		Overall F	DF	Significance	Eta. Squared
	Selection	0.76	(1, 89)	0.764	0.001
	Source	1.09	(1, 89)	0.299	0.012
	Selection * OBSE (H <sub>o2</sub> )	0.88	(1, 89)	0.880	0.000
	Source * OBSE	2.41	(1, 89)	0.124	0.026
	Source * Selection	0.41	(1, 89)	0.407	0.008
	Source * Selection * OBSE	0.50	(1, 89)	0.502	0.001

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*(H<sub>o2</sub>=hypothesis2)*

**Table 11: Length of Time Repeated Measures Results for the  
Developmental Condition**

		Overall F	DF	Significance	Eta. Squared
	Selection	0.36	(1, 89)	0.551	0.004
	Source	1.09	(1, 89)	0.299	0.012
	Selection * OBSE (H <sub>o2</sub> )	0.53	(1, 89)	0.469	0.006
	Source * OBSE	2.41	(1, 89)	0.124	0.026
	Source * Selection	0.00	(1, 89)	0.982	0.000
	Source * Selection * OBSE	0.03	(1, 89)	0.873	0.000

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*(H<sub>o2</sub>=hypothesis2)*

### **Main Analyses: Tests of the hypotheses (Individual Data)**

In this section, a different approach was used to test the hypotheses. In this approach, rather than calculating the mean friendship and task acquaintanceship score for selected versus not selected raters, each selected and not selected rater was treated independently. Therefore, instead of having 2 levels (mean selected and mean not selected), the selection variable now had 4 levels (1<sup>st</sup> selected, 2<sup>nd</sup> selected, not selected A, and not selected B). The data was treated in this manner because it was assumed that when respondents were asked to select 2 raters out of the 4 they had described, the order

with which they made this selection mattered. In other words, it was assumed that the first rater they selected was different from the second. It is plausible that the first rater selected is most representative of whom they would prefer to rate their behavior and the second rater the next representative. Analyzing the data in this manner allowed us to test this distinction and allowed us to test for the difference between the first selected rater and the second, which may be significantly different and which gets lost when averaging the selected versus not selected raters. Note that, while it was possible to distinguish between who was selected first and who was selected second, the two not selected raters could not be ranked as they simply remained unselected. This is an important distinction, as we have no way of knowing which rater they wanted more or less than the other. Therefore, the two not selected raters will simply be labeled not selected A and not selected B.

This breakdown yielded a total of 32 values (16 for friendship and 16 for task acquaintanceship), double the amount in the initial analyses. The breakdown of these variables is illustrated below in Tables 12 and 13.

**Table 12: Friendship Variables (Individual data)**

FRIENDSHIP VARIABLES								
	Administrative				Development			
	1 <sup>st</sup> Selected	2 <sup>nd</sup> Selected	Not Selected (A)	Not Selected (B)	1 <sup>st</sup> Selected	2 <sup>nd</sup> Selected	Not Selected (A)	Not Selected (B)
Peer	FA1P	FA2P	FAAP	FABP	FD1P	FD2P	FDAP	FDBP
Subordinate	FA1S	FA2S	FAAS	FABS	FD1S	FD2S	FDAS	FDBS

**Table 13: Task Acquaintanceship Variables (Individual data)**

TASK ACQUAINTANCESHIP VARIABLES								
	Administrative				Development			
	1 <sup>st</sup> Selected	2 <sup>nd</sup> Selected	Not Selected (A)	Not Selected (B)	1 <sup>st</sup> Selected	2 <sup>nd</sup> Selected	Not Selected (A)	Not Selected (B)
Peer	TA1P	TA2P	TAAP	TABP	TD1P	TD2P	TDAP	TDBP
Subordinate	TA1S	TA2S	TAAS	TABS	TD1S	TD2S	TDAS	TDBS

The variable FASP in Table 2, for example, now becomes FA1P and FA2P (Table 12). FA1P refers to the friendship score(F) when the ratings were to be used for administrative purposes(A) for the first selected(1) peer(P). FA2P refers to the friendship score(F) when the ratings were to be used for administrative purposes(A) for the second selected(2) peer(P). Likewise, FANP (Table 2) now becomes FAAP and FABP (Table 13). FAAP refers to the friendship score(F) when the ratings were to be used for administrative purposes(A) for not selected peer(P) A(A). FABP refers to the friendship score(F) when the ratings were to be used for administrative purposes(A) for not selected peer(P) B(B).

Therefore, although the breakdown of the variables differed in this second approach to data analysis, the analyses performed were the same. The mean friendship and task acquaintanceship scores for both the selected AND the not selected raters was calculated so that we could compare those raters who were selected against those who were not. Therefore, the data was again analyzed separately for each purpose condition (administrative and developmental) since the overall friendship and task acquaintanceship means of each condition would be equal and perfectly correlated (see

Tables 14 & 15 below).

**Table 14: Mean Friendship Score for each condition (Individual data)**

PURPOSE = Administration (mean = 3.761)							
SOURCE							
PEER (Mean = 3.809)				SUBORDINATE (Mean =3.713)			
1 <sup>st</sup> Selected (Mean = 4.143)	2 <sup>nd</sup> Selected (Mean = 3.901)	Not Selected (A) (Mean = 3.621)	Not Selected (B) (Mean = 3.571)	1 <sup>st</sup> Selected (Mean = 3.945)	2 <sup>nd</sup> Selected (Mean = 3.813)	Not Selected (A) (Mean = 3.588)	Not Selected (B) (Mean = 3.506)
PURPOSE = Developmental (mean = 3.761)							
SOURCE							
PEER (Mean =3.809)				SUBORDINATE (Mean =3.713)			
1 <sup>st</sup> Selected (Mean = 4.000)	2 <sup>nd</sup> Selected (Mean = 4.039)	Not Selected (A) (Mean = 3.582)	Not Selected (B) (Mean = 3.615)	1 <sup>st</sup> Selected (Mean = 3.813)	2 <sup>nd</sup> Selected (Mean = 3.835)	Not Selected (A) (Mean = 3.550)	Not Selected (B) (Mean = 3.654)

**Table 15: Mean Task Acquaintanceship Score for each condition (Individual data)**

PURPOSE = Administration (mean =3.806)							
SOURCE							
PEER (Mean =3.870)				SUBORDINATE (Mean =3.742)			
1 <sup>st</sup> Selected (Mean = 4.024)	2 <sup>nd</sup> Selected (Mean = 3.712)	Not Selected (A) (Mean = 3.901)	Not Selected (B) (Mean = 3.841)	1 <sup>st</sup> Selected (Mean = 3.862)	2 <sup>nd</sup> Selected (Mean = 3.859)	Not Selected (A) (Mean = 3.591)	Not Selected (B) (Mean = 3.655)
PURPOSE = Developmental (mean = 3.806)							
SOURCE							
PEER (Mean =3.870)				SUBORDINATE (Mean =3.742)			
1 <sup>st</sup> Selected (Mean = 4.068)	2 <sup>nd</sup> Selected (Mean = 3.846)	Not Selected (A) (Mean = 3.776)	Not Selected (B) (Mean = 3.789)	1 <sup>st</sup> Selected (Mean = 3.798)	2 <sup>nd</sup> Selected (Mean = 3.833)	Not Selected (A) (Mean = 3.675)	Not Selected (B) (Mean = 3.662)

## Friendship

The same analyses performed in the previous friendship section were repeated here. From Table 16 below we see that the main effect for selection was significant [ $F(3, 87) = 3.92, p < 0.01$ ] for the administrative condition. However, the main effect for selection was not significant [ $F(3, 87) = 1.60, p = ns$ ] for the developmental condition (Table 17).

**Table 16: Repeated Measures Results for the Administrative Condition; Friendship Outcome (Individual Data)**

		Overall F	DF	Significance	Eta. Squared
	Selection	3.92	(3,87)	0.011***	0.119
	Source	5.14	(1,89)	0.026**	0.055
	Selection * OBSE ( $H_{01}$ )	2.32	(3,87)	0.081*	0.074
	Source * OBSE	4.44	(1,89)	0.038**	0.047
	Source * Selection	0.06	(3,87)	0.983	0.002
	Source * Selection * OBSE	0.07	(3,87)	0.976	0.002

\*Significant at the 0.10 level

\*\* Significant at the 0.05 level

\*\*\* Significant at the 0.01 level

( $H_{01}$ =hypothesis 1)

**Table 17: Repeated Measures Results for the Developmental Condition; Friendship Outcome (Individual Data)**

		Overall F	DF	Significance	Eta. Squared
	Selection	1.60	(3,87)	0.195	0.052
	Source	5.14	(1,89)	0.026**	0.055
	Selection * OBSE ( $H_{01}$ )	1.03	(3,87)	0.385	0.034
	Source * OBSE	4.44	(1,89)	0.038**	0.047
	Source * Selection	0.47	(3,87)	0.705	0.016
	Source * Selection * OBSE	0.35	(3,87)	0.787	0.012

\*Significant at the 0.10 level

\*\* Significant at the 0.05 level

\*\*\* Significant at the 0.01 level

( $H_{01}$ =hypothesis 1)

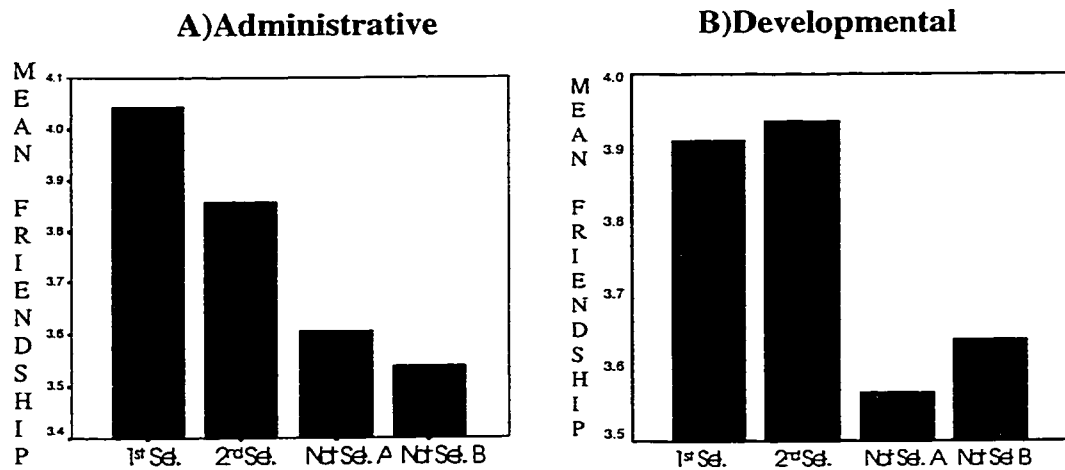
Therefore, hypothesis 3, which stated that when the purpose of the appraisal was administrative, individuals would be more likely to choose raters who they



perceived as being higher on friendship, was supported. In order to determine exactly where the differences lie, a series of post-hoc tests were performed. Since these were post-hoc tests, a more stringent p-value (0.01) was used. The results revealed that for the administrative condition, the friendship score of the first person selected differed significantly from not selected person A [ $t(90)=7.82, p<0.01$ ] and not selected person B [ $t(90)=8.09, p<0.01$ ]. In addition, the friendship score of the second person selected differed significantly from not selected person A [ $t(90)=3.38, p<0.01$ ] and not selected person B [ $t(90)=3.98, p<0.01$ ]. The friendship score of the first and second person selected differed significantly [ $t(90)=2.72, p<0.01$ ]. However, the friendship score of the two not selected individuals [ $t(90)= -0.93, p=ns$ ] did not differ significantly.

For the developmental condition, the friendship score of the first person selected differed significantly from not selected person A [ $t(90)=4.19, p<0.01$ ] and not selected person B [ $t(90)=3.26, p<0.01$ ]. In addition, the friendship score of the second person selected differed significantly from not selected person A [ $t(90)=5.22, p<0.01$ ] and not selected person B [ $t(90)=3.69, p<0.01$ ]. However, the friendship score of the first and second person selected [ $t(90)= -0.43, p=ns$ ] and the two not selected individuals [ $t(90)= -0.98, p=ns$ ] did not differ significantly. These relationships are displayed graphically in Figure 3 below.

**Figure 3: Selection by Purpose interaction; friendship (Individual data)**



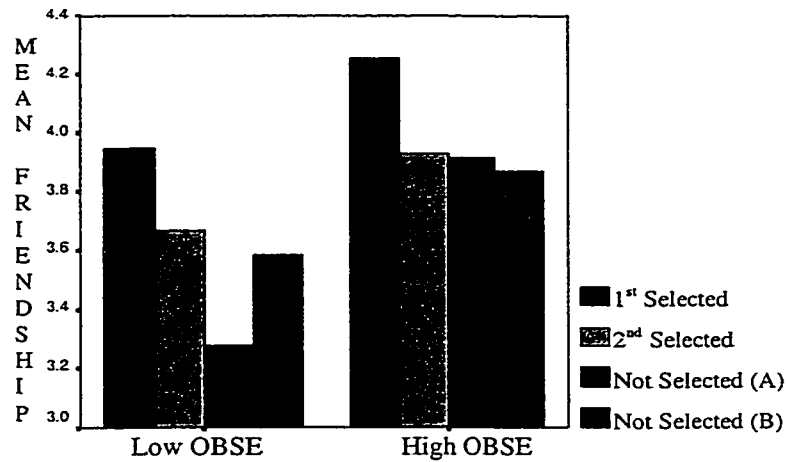
From the figure we can see that the mean friendship incrementally drops in the administrative condition, with the first selected rater being higher in perceived friendship than the second selected rater. This pattern does not repeat itself for the developmental condition however. In addition, the mean difference in perceived friendship between the first selected rater and the not selected raters is larger in the administrative condition than in the developmental condition. Therefore, when the purpose of the appraisal is administrative, respondents seem to differentiate raters based on friendship more so than when the appraisal is to be used for developmental purposes.

In terms of the main effect for source, and the source by OBSE interaction, the results are the same here (Tables 16 & 17) as they were in the first approach to data analysis (Tables 6 & 7). Again, this is because the main effect for source and the source by OBSE interaction analyses do not include the selection manipulation. These analyses simply compare the mean friendship score for peers (both those selected and those not selected) and subordinates (both those selected and those not selected). Therefore, they

cannot differ when the analyses are analyzed in terms of means (selected / mean not selected) versus individually (1<sup>st</sup> selected / 2<sup>nd</sup> selected / not selected A and B).

The test for hypothesis 1 can be found by examining the selection by OBSE interaction. As demonstrated in Tables 16 and 17 above, the selection by OBSE interaction was marginally significant when the purpose of the appraisal was administrative [ $F(3, 87) = 2.32, p < 0.10$ ] and was not significant when the purpose of the appraisal was developmental [ $F(3, 87) = 1.03, p = ns$ ]. Therefore, hypothesis 1, which stated that individuals low in OBSE would be more likely than individuals high in OBSE to choose raters who they perceived as being higher in friendship to rate their behavior at work was supported in the administrative condition. This relationship is illustrated in Figure 4 below.

**Figure 4: Selection by OBSE interaction for friendship:  
Administrative condition (Individual data)**



From the figure, we see that the difference in friendship between selected versus not selected raters is larger for individuals low in OBSE than for those high in OBSE. Therefore, individuals low in OBSE tend to make more of an effort to select those raters who are their friends when the purpose of the appraisal is administrative. This pattern is in the same direction as hypothesis 1.

In terms of the exploratory analyses involving rater source, the source by selection interaction was not significant for both the administrative purpose [ $F(3, 87) = 0.06, p=ns$ ] and the developmental purpose [ $F(3, 87) = 0.47, p=ns$ ] (Tables 16 & 17). Respondents do not select peers versus subordinates differently depending on their perceived level of friendship. Therefore, although the main effect for source illustrates that peers were perceived as being better friends than subordinates, this did not seem to determine the selection of raters.

The source by selection by OBSE interaction was also not significant for both the administrative purpose [ $F(3,87)=0.07, p=ns$ ] and the developmental purpose [ $F(3,87)=0.35, p=ns$ ] (Tables 16 & 17). Therefore, respondent's level of OBSE did not influence the selection of peers versus subordinates depending on their level of perceived friendship. Overall these results suggest that individuals do not choose peers versus subordinates differently based on the level of friendship they have with them. In addition, OBSE plays no role in influencing this relationship.

#### Task Acquaintanceship

The same analyses performed in the first approach to data analysis was repeated here. From Table 18 below we see that the main effect for selection was not significant [ $F(3, 87) = 1.89, p=ns$ ] for the administrative condition. However, the main effect for selection was marginally significant [ $F(3, 87) = 2.56, p<0.10$ ] for the developmental condition (Table 19).

**Table 18: Repeated Measures Results for the Administrative Condition:  
Task Acquaintanceship Outcome (Individual Data)**

		Overall F	DF	Significance	Eta. Squared
	Selection	1.89	(3,87)	0.137	0.061
	Source	1.50	(1,89)	0.224	0.017
	Selection * OBSE ( $H_{o2}$ )	1.28	(3,87)	0.286	0.042
	Source * OBSE	0.80	(1,89)	0.375	0.009
	Source * Selection	0.23	(3,87)	0.874	0.008
	Source * Selection * OBSE	0.19	(3,87)	0.905	0.006

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*( $H_{o2}$ =hypothesis 2)*

**Table 19: Repeated Measures Results for the Developmental Condition:  
Task Acquaintanceship Outcome (Individual Data)**

		Overall F	DF	Significance	Eta. Squared
	Selection	2.56	(3,87)	0.060*	0.081
	Source	1.50	(1,89)	0.224	0.017
	Selection * OBSE ( $H_{o2}$ )	2.08	(3,87)	0.109*	0.067
	Source * OBSE	0.80	(1,89)	0.375	0.009
	Source * Selection	0.53	(3,87)	0.662	0.018
	Source * Selection * OBSE	0.41	(3,87)	0.750	0.014

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

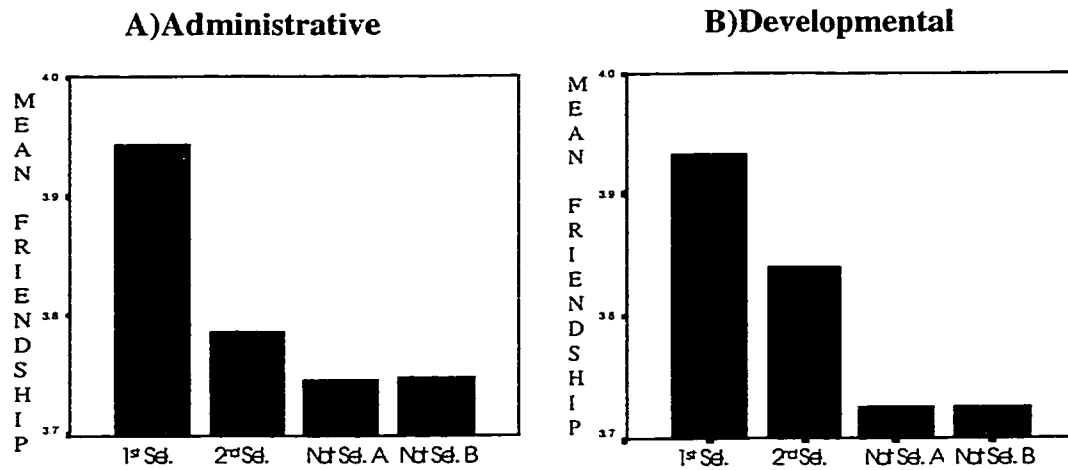
*( $H_{o2}$ =hypothesis 2)*

This means that when the purpose of the appraisal was developmental, respondent's were more likely to select raters who were high in task acquaintanceship. This corresponds to hypothesis 3, which stated that when the purpose of the appraisal was administrative, individuals would be more likely to choose raters who they perceived as being higher on friendship. The contrary is that when the purpose of the appraisal is developmental, individuals would be more likely to select raters who are high in task acquaintanceship. The results support this contention.

In order to determine exactly where the differences lie, paired sample t-tests were performed. Since these were post-hoc tests, a more stringent p-value (0.01) was used. The results revealed that for the administrative condition, the task acquaintanceship score of the first person selected differed significantly from not selected person A [ $t(90)=3.38, p<0.01$ ] and not selected person B [ $t(90)=3.00, p<0.01$ ]. However, the task acquaintanceship score of the second person selected did not differ significantly from not selected person A [ $t(90)=0.59, p=ns$ ] and not selected person B [ $t(90)=0.46, p=ns$ ]. In addition, the task acquaintanceship score of the first and second person selected differed significantly [ $t(90)=2.44, p<0.01$ ]. However, the task acquaintanceship score of the two not selected individuals [ $t(90)= -0.93, p=ns$ ] did not differ significantly.

For the developmental condition, the task acquaintanceship score of the first person selected differed significantly from not selected person A [ $t(90)=2.85, p<0.01$ ] and not selected person B [ $t(90)=2.80, p<0.01$ ]. However, the task acquaintanceship score of the second person selected did not differ significantly from not selected person A [ $t(90)=1.66, p=ns$ ] and not selected person B [ $t(90)=1.43, p=ns$ ]. The task acquaintanceship score of the first and second person selected [ $t(90)= 1.49, p=ns$ ] and the two not selected individuals [ $t(90)= -0.00, p=ns$ ] did not differ significantly. These relationships are displayed graphically in Figure 5 below.

**Figure 5: Selection by Purpose interaction; task acquaintanceship  
(Individual data)**



From the figure we can see that the mean difference in perceived task acquaintanceship between the first selected rater and the not selected raters is larger in the developmental condition than in the administrative condition. Therefore, when the purpose of the appraisal is developmental, respondents seem to differentiate raters based on task acquaintanceship more so than when the appraisal is to be used for developmental purposes.

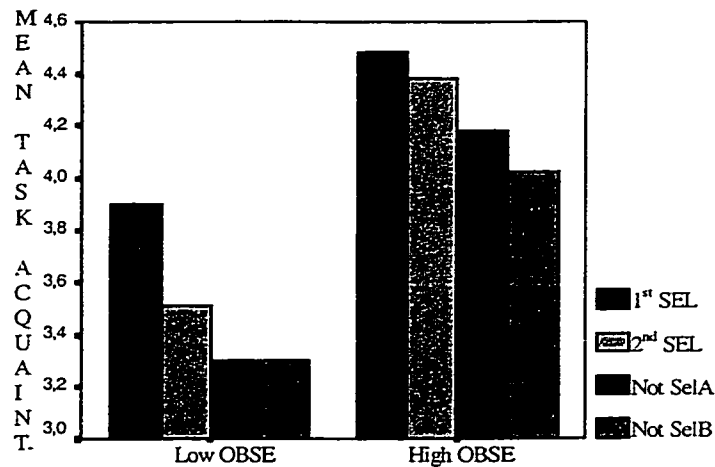
Again, in terms of the main effect for source, and the source by OBSE interaction, the results are the same here (Tables 18 & 19) as they were in the first approach to data analysis (Tables 8 & 9).

The test for hypothesis 2 can be found by examining the selection by OBSE interaction. As demonstrated in Tables 18 and 19 above, the selection by OBSE interaction was not significant when the purpose of the appraisal was administrative [ $F$



(3, 87) = 1.28,  $p = ns$ ], however, it was marginally significant when the purpose of the appraisal was developmental [ $F(3, 87) = 2.08, p < 0.10$ ]. Therefore, hypothesis 2, which stated that ratees high in OBSE would be more likely than ratees low in OBSE to choose raters who they perceived as having a high degree of task acquaintanceship with their work was not supported when the purpose of the appraisal was administrative. It was supported when the purpose of the appraisal was developmental however. This relationship is illustrated in Figure 6 below.

**Figure 6: Selection by OBSE interaction for task acquaintanceship;  
Developmental condition (Individual data)**



From the figure, we see that the difference in task acquaintanceship between selected versus not selected raters is larger for individuals low in OBSE than for those high in OBSE. Therefore, individuals low in OBSE tend to make more of an effort, than those individuals who are high in OBSE to select those raters who are high in task acquaintanceship when the purpose of the appraisal is developmental. This pattern is in the opposite direction as hypothesis 2.

In terms of the exploratory analyses involving rater source, the source by selection interaction was not significant for both the administrative purpose [ $F(3, 87) = 0.23, p=ns$ ] and the developmental purpose [ $F(3, 87) = 0.53, p=ns$ ] (Tables 18 & 19). Respondents did not select peers versus subordinates differently depending on their perceived level of task acquaintanceship.

The source by selection by OBSE interaction was also not significant for both the administrative purpose [ $F(3,87)=0.19, p=ns$ ] and the developmental purpose [ $F(3,87)=0.41, p=ns$ ] (Tables 8 & 9). Therefore, respondent's level of OBSE did not influence the selection of peers versus subordinates depending on their level of perceived task acquaintanceship.

#### Length of time

As we can see in Tables 20 and 21 below, no main effect for selection was found for the administrative condition [ $F(3,87)=0.10, p=ns$ ] or the developmental condition [ $F(3,87)=0.17, p=ns$ ]. In addition, the selection by OBSE interaction was not significant for the administrative condition [ $F(3,87)=0.10, p=ns$ ] or the developmental condition [ $F(3,87)=0.30, p=ns$ ]. Therefore, hypothesis 2 was not supported.

**Table 20: Length of Time Repeated Measures Results for the  
Administrative Condition (Individual Data)**

		Overall F	DF	Significance	Eta. Squared
	Selection	0.10	(3,87)	0.834	0.010
	Source	1.09	(1,89)	0.299	0.012
	Selection * OBSE ( $H_{02}$ )	0.10	(3,87)	0.962	0.003
	Source * OBSE	2.41	(1,89)	0.124	0.026
	Source * Selection	0.40	(3,87)	0.752	0.014
	Source * Selection * OBSE	0.42	(3,87)	0.719	0.015

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*( $H_{02}$ =hypothesis2)*

**Table 21: Length of Time Repeated Measures Results for the  
Developmental Condition (Individual Data)**

		Overall F	DF	Significance	Eta. Squared
	Selection	0.17	(3,87)	0.878	0.008
	Source	1.09	(1,89)	0.299	0.012
	Selection * OBSE ( $H_{02}$ )	0.30	(3,87)	0.707	0.016
	Source * OBSE	2.41	(1,89)	0.124	0.026
	Source * Selection	0.24	(3,87)	0.671	0.018
	Source * Selection * OBSE	0.25	(3,87)	0.580	0.022

*\*Significant at the 0.10 level*

*\*\* Significant at the 0.05 level*

*\*\*\* Significant at the 0.01 level*

*( $H_{02}$ =hypothesis2)*

## **DISCUSSION**

### **Review of Results**

The data collected in this study was analyzed using two different approaches. The first involved treating the selected raters as one group and the not selected raters as the second group. Although this approach was useful in testing the hypotheses, aggregating the two selected raters into one group meant that the difference between the first and second selected rater would be unable to be determined. Therefore, if individuals chose their raters in a ranking order with the first rater being the most representative of whom they would have liked to evaluate their behavior and the second rater being the next representative, summing their scores would not allow us to examine the effects of each independently.

As a result, a second approach to data analysis was conducted. In this approach, the friendship and task acquaintanceship scores of the first selected rater, the second selected rater and the two not selected raters (A and B) were kept separate. Therefore, the selection variable had four levels rather than two. Since we used a web-based survey, we knew that the first selected rater was the first to be entered in the data bank and the second selected rater was the second to be entered in the data bank. The two not selected raters however could not be ranked as they simply remained unselected.

Based on the first approach to data analysis, we can conclude that respondents rated their peers as being higher in perceived friendship than their subordinates. This

makes logical sense given the different power dynamics that exist at work. It is common practice for individuals to become better friends with their peers, rather than their subordinates. In addition, the results revealed that the source by OBSE interaction was significant. Individuals low in OBSE seemed to make a greater differentiation in the level of friendship they had with their subordinates versus their peers. Individuals high in OBSE did not make such a big differentiation. From this we can conclude that maybe those individuals who are lower in OBSE do not interact with their subordinates in a friendly manner and do not try to build a friendly relationship with them. While these results are interesting, they are not the main focus of the current study.

Based on the results obtained from the second approach to data analysis, we can conclude that individuals do select their potential raters differently depending on how good of a friend and how task acquainted they perceive them to be. This was evidenced by the main effect of selection for both friendship (administrative condition) and task acquaintanceship (developmental condition). Therefore, this study has identified two of the dimensions along which individuals differentiate raters. While this study did not test for the different ratings obtained when individuals choose friends over individuals who are highly task acquainted with their work, it did examine and find that ratees differentiated amongst raters based on these two dimensions. Individuals will select raters who they are friends with and who are acquainted with their tasks at work more so than those who are not their friends and are not acquainted with their tasks. The practical implications of this finding is that raters do have a hand in influencing their performance

evaluations when they are given the opportunity to choose their raters. They do not select potential raters randomly.

Therefore, although MSF was initially introduced in order to provide a more reliable and valid assessment of work performance, this study seems to suggest that allowing ratees to select their raters may actually be undermining this outcome. New errors and biases may be introducing themselves into the process. This finding is especially important given that Brutus and Derayeh (2000) in their study on 360-degree practices, found that 84% of the companies interviewed gave their employees complete freedom in choosing their raters.

### Hypotheses

The current study hypothesized that an individual's level of OBSE may play a role in influencing which raters were selected. Specifically, hypothesis 1 stated that individuals low in OBSE would be more likely than individuals high in OBSE to choose raters who they perceive as being high in friendship to rate their behavior at work. This hypothesis was supported by the second approach to data analysis, in the administrative condition. Therefore, respondents in this study who were low in OBSE were more likely than individuals high in OBSE to select those raters who were high on friendship when the purpose of the appraisal was administrative. As we have assumed throughout this paper, this was done in an effort to obtain a positive evaluation of their work performance. This finding is in line with the theory of self-enhancement discussed

earlier (Dipboye, 1977). The theory proposes that individuals have a need to view themselves as favorably as possible (simple self-enhancement). But, the strength of the need for self-enhancement is directly related to the extent to which this need has been thwarted in the past. Therefore, individuals low in self-esteem would be more likely to strive toward self-enhancement than their high self-esteem counterparts (compensatory self-enhancement). This finding is especially important since as discussed earlier, it illustrates that not all individuals will select raters who will rate them positively in an effort to enhance their evaluations. Therefore, some evaluations may be enhanced and inaccurate and others may not, yielding biased performance evaluations.

Consequently, in a multi-source feedback process in which employees are given the opportunity to choose who will rate their performance, persons with a low level of OBSE may be more likely to seek feedback from their friends, which we are assuming will be overly positive. Individuals low in OBSE see a potentially negative performance evaluation as more threatening to their self-concept. Therefore, they engage in impression management (Schlenker, 1980) and choose those raters who they believe are most likely to rate them positively, their friends.

However, the practical implications of avoiding negative performance information in the work setting can be very costly. Employees need to know what they are doing wrong and the areas of their work that need improvement if they are to grow and develop themselves (Ashford & Tsui, 1991). If employees are not made aware of their weaknesses, then they cannot take the steps to correct or improve their behavior.

This suggests that those individuals who opted for raters who were high on task acquaintanceship, received more accurate ratings and will be better off in terms of long-term development and growth. However, the results from this study did not support the contention that individuals high in OBSE would be more likely to select individuals high in task acquaintanceship.

Specifically, Hypothesis 2, stated that the selection of raters according to perceived task acquaintanceship would be moderated by OBSE in that individuals high in OBSE would be more likely than individuals low in OBSE to choose raters who have a high degree of task acquaintance with their work. Although the theory of self-assessment and the research which finds that individuals who are high in OBSE are most likely to seek out self-diagnostic and accurate information about themselves (Troe, 1975) supports this hypothesis, the results in this study did not. Rather, the results from this study marginally supported the contention that individuals low in OBSE make more of an effort than those individuals who are high in OBSE, to select raters whom are high in task acquaintanceship when the purpose of the appraisal is developmental. This pattern is in the opposite direction as hypothesis 2, and warrants further investigation.

Individuals low in OBSE are those individuals who do not feel like their contributions at work are satisfactory. Therefore, the positive evaluations that they intentionally seek out about their work performance may be unwarranted and inaccurate. This group of individuals may be especially in need of accurate performance assessments which detail areas of improvement so that they can ameliorate their



behavior. Selectively seeking out positive feedback may allow them to feel good about themselves in the short-term but the long-term implications will be stunted development and improvement in the area of work. This will act to reinforce the feelings they hold about their unsatisfactory work performance and consequently their low level of OBSE will be maintained. In essence a vicious circle may be created in which individuals low in OBSE, through their influence in controlling the formal performance feedback which they receive at work, will perpetuate their level of low OBSE. However, the results do not suggest that this pattern of behavior is in operation. In fact individuals low in OBSE may be more likely than individuals high in OBSE to be seeking accurate assessments when the evaluation is to be used for developmental purposes.

Hypothesis 3 stated that regardless of an individual's level of OBSE, when feedback will be used for administrative purposes, individuals will be more likely to seek raters who are high in friendship than when feedback will be used for developmental purposes. In this study, the second approach to data analysis revealed that individuals selected raters who were higher on friendship when the purpose of the appraisal was administrative. There are two reasons why this may operate. First, there is more to be gained by choosing friends who will most likely give a positive performance evaluation when the appraisal influence decisions concerning compensation, promotion, terminations or lateral transfers. Secondly, when the performance evaluations are to be administrative and therefore public, research supports the fact that individuals are more likely to engage in defensive impression management and 'face-saving' (Morrison and Bies, 1991). This finding is especially interesting

because in practice, it is often the case that ratees are given full freedom in choosing their raters when the purpose of the appraisal is developmental. However, when the purpose of the appraisal is administrative, their selection is often monitored in some manner. The results from this study support this practice.

In addition, the results also revealed that when the purpose of the appraisal was developmental, raters who were higher in task acquaintanceship were most likely to be selected. Therefore, when the purpose of the appraisal was developmental, individuals chose those peers that were higher in task acquaintanceship. Given that the appraisal information would remain private and only be utilized for personal development, individuals may have reasoned that those individuals that are highest in task acquaintanceship would provide the most accurate performance assessment.

Lastly, rater source was included in this study as an exploratory variable. Given the different work and power dynamics that exist between peers and subordinates, it was suggested in this paper that ratees may use different strategies when choosing amongst these different rating sources. However, no significant results were found. Therefore, ratees did not seem to use different strategies when they were selecting peers versus subordinates. Although, Ashford and Tsui (1991) found that managers would seek more negative (and thereby more accurate) feedback from subordinates and the least amount of negative feedback from peers, these results were not replicated in this study.

## **Limitations**

One of the limitations of this study was the small sample size ( $n=91$ ). Given the statistical fact that a smaller sample size leads to less power, a larger sample would have been preferable. A larger sample would have increased power and made the probability of detecting effects much greater. Nonetheless, since statistically significant results were found in this study, the strength of the interactions must have been quite salient.

A second limitation was the potential of response or sampling bias. The target sample was MBA students in eight major universities across Canada who worked full-time and had a minimum of four subordinates under their supervision. This group of individuals may be different in many ways from the general population. In addition, those individuals who did take the time to complete the survey may have differed still. Nonetheless, a look at the demographics of our sample indicates that our sample was quite diverse.

A third limitation was the artificial or simulated nature of the research procedure. Although respondents were asked to describe and select four actual peers and subordinates, the MSF process was fictitiously simulated and not gathered in an organization that had the process in place. Therefore, the generalizability and external validity of the results presented in this paper remain to be determined.

A final limitation was the negatively skewed OBSE variable. This means that most individuals rated themselves as being high on the OBSE dimension. This may have been expected since our sample was MBA students. However, skewed data violates the assumption of normality for multivariate statistics (Tabachnick & Fidell, 1996). Although data transformations were conducted, they had little effect in normalizing the data.

### **Directions for future research**

While this study examined what influence OBSE may have in the selection of raters, there are many other individual differences that may influence this process. For example, individuals who have an internal locus of control may be especially likely to selectively choose raters because they believe they have the capacity to influence the results. In addition, organizational culture may influence the rater selection process. For example, in organizations that have a very strong growth and development culture, actively seeking only positive evaluations may be unthinkable.

Another direction for future research involves empirically testing the assumption that choosing friends does in fact lead to enhanced performance evaluations. Likewise, the assumption that choosing individuals high on task acquaintanceship does in fact lead to more accurate and potentially negative performance evaluations should also be empirically tested. These assumptions formed the basis of the current thesis and while there is much research that suggests that this relationship does in fact occur (as reviewed

in the introduction of this paper), no research has empirically tested for this effect in a MSF context. If these assumptions would be disconfirmed, then selectively choosing friends in a MSF process would not lead to biases or render the performance evaluation as less valid.

In addition, investigating the effects of *not* having ratees select their raters is also worth exploring. Although the data presented in this paper suggests that allowing ratees to select their raters may lead to biases in the resulting performance evaluation data, not involving ratees in the process may lead to decreased acceptance and ownership of the process. If this were to occur, individuals would be less interested in making use of the results.

Although the effects that an inaccurate performance evaluation may have on the ratee was discussed in this study, no research exists which actually measures or attempts to quantify in some manner the long-term effects that biased performance information may have. For example, to what extent does it lead to stunted individual and organizational growth? In addition, do inaccurate performance evaluations alter important personnel decisions such as retention, promotion, and salary increases?

Lastly, other ways in which ratees may influence their performance evaluation data should also be explored. Rater selection may just be one way in which ratees may influence the information they receive. Other forms of manipulation which are worth exploring involve: the subtle communication of benefits which can be gained if

subordinates rate their boss' performance positively; the agreement amongst employees that they will both rate each other favorably; the distribution of evaluation forms just after the delivery of good news etc..

## **CONCLUSION**

The practical implications of these findings are many and will be summarized in this section. The results from this study support the contention that when ratees are allowed to select their raters, they do not act passively. Rather, they do select raters differently depending on the level of friendship they have with that individual and how task acquainted they are with their work. In addition, the ratees' level of OBSE influences the process. Therefore, while MSF was introduced as a method aimed at decreasing biases and increasing the validity and reliability of performance evaluation data, new biases may be introduced when ratees are allowed to select their raters. As a consequence, inaccurate and distorted performance information is gathered and the developmental and growth outcomes that accurate feedback is supposed to yield are compromised.

So where do we go from here? Should the ratee be devoid of all opportunity to select their raters or should some compromise be reached? I would suggest a compromise. Allowing the ratee to be involved in the selection of their raters gives them the opportunity to be involved in the process. It allows them to commit to the process and to feel like they had a hand in creating it rather than having it imposed upon them. With this, they can take ownership of the results and give credit to the process. Therefore, perhaps the best solution to this dilemma would be to allow ratees to select their raters and then have this selection validated by their immediate supervisor. In addition, employees should be educated on the long-term growth and developmental benefits that can be gained from an accurate performance evaluation.

## **References**

- Alexander, E.R. & Wilkins, R.D. (1982). Performance Rating Validity: The Relationship of Objective and Subjective Measures of Performance. *Group & Organization Studies*, Vol.7, No.4. 485- 497.
- Alloy, L.B. & Abramson, L.Y. (1988). Depressive realism: Four theoretical perspectives. In L.B. Alloy (Ed.), *Cognitive processes in depression* (223-265). New York: Guilford Press.
- Ammons, R.B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. *Journal of General Psychology*, Vol.54, 279-299.
- Antonioni, D. (1994). The effects of feedback accountability on upward appraisal ratings. *Personnel Psychology*, Vol.47, 349-356.
- Atwater, L. & Waldman, D. (1998). Accountability in 360-degree feedback. *HR Magazine*, Vol.43, No.6, 96-104.
- Ashford, S.J. & Cummings, L.L. (1983). Feedback as an individual resource: Personal strategies of creating information. *Organizational Behavior and Human Performance*, Vol.32, 370-398.
- Ashford, S.J. & Northcraft, G.B.(1992). Conveying more (or less) than we realize: The role of impression-management in feedback-seeking. *Organizational Behavior And Human Decision Processes*, Vol. 53, No. 3, 310-335.
- Ashford, S.J. (1986). Feedback-seeking in individual adaptation: A resource perspective. *Academy of Management Journal*, Vol. 29, 465-487.
- Ashford, S.J. & Tsui, A.S. (1991). Self-regulation for managerial effectiveness: The role of active feedback seeking. *Academy of Management Journal*, Vol.34, No.2, 251-280.
- Balzer, W.K., & Sulsky, L.M. (1992). Halo and performance appraisal: A critical examination. *Journal of Applied Psychology*, Vol. 77, No.6, 975-985.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Reviews*, Vol.84, 191-215.
- Baumeister, R.F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, Vol.91, 3-26.
- Becker, T.E. & Klimoski, R.J. (1989). A field study of the relationship between the organizational feedback environment and performance. *Personnel Psychology*, Vol.42, 343-358.



Belcourt, M., Sherman, A.W., Bohlander, G.W. & Snell, S. (1996). *Managing Human Resources*, International Thomson Publishing, Scarborough, Ontario.

Bernardin, H.J., & Beatty, R.W. (1984). *Performance Appraisal: Assessing human Behavior at work*, Boston, Kent Publishing Company.

Bettenhausen, K.L. & Fedor, D.B. (1997). Peer and Upward Appraisals: A comparison of their benefits and their problems. *Group & Organization Management*, Vol.22, No.2, 236-263.

Blascovich, J. & Tomaka, J. (1991). Measures of self-esteem, in *Measures of personality and social psychological attitudes. Measures of social psychological attitudes*. Robinson, John P. (Ed); Shaver, Phillip R. (Ed) Vol. 1. (115-160). San Diego, CA.

Bohl, D.L. (1996). Minisurvey: 360-degree appraisals yield superior results, survey shows. *Compensation and Benefits Review*, Vol.28, No.5, 16-19.

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

Bradley, G.W. (1978). Self-serving biases in the attribution processes: A reexamination of the fact or fiction question. *Journal of Personality and Social Psychology*, Vol.36, 56-71.

Brickman, P., & Bullman, R.J. (1977). Pleasure and pain in social comparison. In J.Suls & R.L.Miller (Eds.), *Social Comparison processes: Theoretical and empirical perspectives*. Washington, D.C.: Halsted:Wiley.

Brockner, J. & Elkind, M. (1985). Self-esteem and reactance: Further evidence of attitudinal and motivational consequences. *Journal of Experimental and Social Psychology*, Vol.21, No.4, 346-361.

Brockner, J. (1988). *Self-esteem at work: Research, theory, and practice*. Lexington, Mass.: D.C. Heath & Co.

Brutus, S. & Derayeh, M. (2000). Current Practices in Multi-Source Assessment. *Unplished Manuscript*.

Cardy, R.L. & Dobbins, G.H. (1986). Affect and Appraisal Accuracy: Liking as an Integral Dimension in Evaluating Performance. *Journal of Applied Psychology*, Vol.71, No.4, 672-679.

Carver, C.S. & Scheier, M.F. (1981). *Attention and Self-regulation: A control theory approach to Human Behavior*. New York: Springer-Verlag.

Carver, C.S. & Ganellen, R.J. (1983). Depression and components of self-punitiveness: High standards, self-criticism and overgeneralization. *Journal of Abnormal Psychology*, Vol. 49, 722-728.

Cascio, W. (1987). *Applied Psychology in Personnel Management* (3<sup>rd</sup> ed.). Englewood Cliffs, NJ: Prentice Hall.

Cederblom, D. & Lounsbury (1980). An investigation of user acceptance of peer evaluation. *Personnel Psychology*, Vol. 42, 37-52.

Chapanis, A. (1964). Knowledge of performance as an incentive in repetitive, monotonous tasks. *Journal of Applied Psychology*, Vol.48, 263-267.

Chappelow, C.T. (1998). 360-degree feedback. In C.D. McCauley, R.S. Moxley, & E. Van Velsor (Eds.), *Center for Creative Leadership handbook of Leadership Development* (29-65). San Francisco, CA: Jossey-Bass.

Church, A.H., Rogelberg, S.G. & Waclawski, J. (2000). Since when is no news good news? The relationship between performance and response rates in multirater feedback. *Personnel Psychology*, Vol.53, 435-451.

Coates, D.E. (1998). Don't tie 360 feedback to pay. *Training*, Vol.35, No.9, 68-78.

Coopersmith, J. (1967). *The antecedents of self-esteem*. San Francisco: Freeman.

Dallessio, A.T. (1998). Using multi-source feedback for employee development and personnel decisions. In Smither J.W. (Ed.), *Performance Appraisal: State of the art in practice* (278-330). San Francisco: Jossey-Bass.

DeNisi, A.S., Cafferty, T.P., & Meglino, B.M. (1984). A Cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behavior and Human Performance*, Vol.33, 360-396.

Dipboye, R.L. (1977). A critical review of Korman's self-consistency theory of work motivation and occupational choice. *Organizational Behavior and Human Performance*, Vol.18, 108-126.

Dipboye, R.L. (1985). Some Neglected Variables in Research on Discrimination in Appraisals. *The Academy of Management Review*, Vol.10, No.1, 116-128.

Dobbins, G.H. & Russell, J.M. (1986). The Biasing Effects of Subordinate Likeableness on Leaders' Responses to Poor Performers: A Laboratory and a Field Study. *Personnel Psychology*, Vol.39, No.4, 759-778.

Dubin, R. (1956). Industrial workers' world: A study of the "central life interests" of industrial workers. *Social Problems*, Vol.3, 131-142.

Eder, R.W. & Fedor, D.B. (1989). Impression management: Its interpretive role in the supervisor-employee feedback process. In R.A. Giacalone & P. Rosenfeld (Eds.), *Impression management in the organization*: 327-342. Hillsdale, NJ: Lawrence Erlbaum Associates.

Epstein, S. (1979). The stability of behavior: 1. On predicting most of the people most of the time. *Journal of Personality and Social Psychology*, Vol.37: 1097-1126.

Farr, J.L. & Newman, D.A. (2000). Rater Selection: Sources of Feedback. In Bracken D., Timmreck, C.W., & Church, A. (Eds.). *Handbook of Multi-source Feedback*. Jossey-Bass: San Francisco.

Fedor, D.B. (1991). Recipient Response to Performance Feedback: A proposed model and its implications. *Research in Personnel and Human Resources Management*, Vol. 9, 73-120.

Ferguson, L.W. (1949). The effect upon appraisal scores of individual differences in the ability of superiors to appraise subordinates. *Personnel Psychology*, Vol. 2, 377-382.

Festinger, L. (1957). *A theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson.

Freeberg, N.E. (1969). Relevance of rater-ratee acquaintance in the validity and reliability of ratings. *Journal of Applied Psychology*, Vol. 53, No.6, 518-524.

French, J.R.P. & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in Social Power* (150-167). Ann Arbor: University of Michigan Press.

Frey, D. & Stahlberg, D. (1986). Selection of information after receiving more or less reliable self-threatening information. *Personality and Social Psychology Bulletin*, Vol 12, No. 4, 434-441.

Frey, D. (1981). The effect of negative feedback about oneself and cost of information on preferences for information about the source of this feedback. *Journal of Experimental Social Psychology*. Vol.17, No.1, 42-50.

Fried, Y., Tiegs, R.B. & Bellamy, A.R. (1992). Personal and interpersonal predictors of supervisors' avoidance of evaluating subordinates. *Journal of Applied Psychology*, Vol.77, 462-468.

Funder, D.C. & Colvin, C.R. (1988). Friends and Strangers: acquaintanceship, agreement and the accuracy of personality judgment. *Journal of Personality and Social Psychology*, Vol.55, 149-158.

Gacalone, R.A. (1989). Image Control: The strategies of impression management. *Personnel*, Vol.6, No.5, 52-55.

Gelfand, D.M. (1962). The influence of self-esteem on rate of verbal conditioning and social matching behavior. *Journal of Abnormal and Social Psychology*, Vol.65, 259-265.

Gilliland, S.W. & Langdon, J.C. (1998). Creating performance management systems that promote perceptions of fairness. In J.W. Smither (Ed.), *Performance Appraisal: State of the art and practice* (209-243). San Francisco: Jossey-Bass.

Greenwald, A.G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, Vol.35, 603-618.

Greller, M.M. & Herold, D.M. (1975). Sources of Feedback: A preliminary Investigation. *Organizational Behavior and Human Performance*, Vol.13, 244-256.

Guion, R.M. (1965). *Personnel Testing*. New York: McGraw-Hill.

Guion, R.M. (1980). On Trinitarian doctrines of validity. *Professional Psychology*, Vol.11, 385-398.

Gupta, N. & Jenkins, D.G. Jr. (1996). The politics of pay. *Compensation and Benefits Review*, March/April Issue, 23-30.

Harris, M.M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, Vol.41, 43-62.

Hedge, J.W. & Kavanagh, M.J. (1988). Improving The Accuracy Of Performance Evaluations. *Journal of Applied Psychology*, Vol.73, No.1, 68-73.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Herold, D. & Greller, M. (1977). Feedback: The definition of a construct. *Academy of Management Journal*, Vol.20, 142-147.

Herold, D.M. & Parsons, C.K. (1985). Assessing the feedback environment in work organizations: Development of the job feedback survey. *Journal of Applied Psychology*, Vol.70, 290-309.

Herold, D.M. & Fedor, D.B. (1998). Individuals' interaction with their feedback environment: The role of domain-specific individual differences. *Research in Personnel and Human Resources Management*, Vol.16, 215-254.

Holzbach, R.L. (1978). Rater Bias in Performance Ratings: Superior, Self and Peer Ratings. *Journal of Applied Psychology*, Vol.63, No.5, 579-592.

- Ilgen, D.R., Fisher, C.D., & Taylor, M.S. (1979). Consequences of individuals feedback on behavior in organizations. *Journal of Applied Psychology*, Vol.64, 359-371.
- Jacobs, R. & Kozlowski, S.W.J. (1985). A Closer Look at Halo Error in Performance Ratings. *Academy of Management Journal*, Vol. 28, No.1, 201-213.
- Jones, E.E. & Gerard, H.B. (1967). *Foundations of Social Psychology*. New York: Wiley.
- Jones, S.C. (1973). Self and interpersonal evaluations: Esteem theories versus consistency theories. *Psychological Bulletin*, Vol.79, 185-199.
- Jones, E.E. & Pittman, T.S. (1982). *Toward a general theory of strategic self-presentation*. In J. Suls (Ed.), *Psychological perspectives on the self* (231-261). Hillsdale, NJ: Erlbaum.
- Karl, K.A. & Kopf, J.M. (1994). Will individuals who need to improve their performance the most, volunteer to receive videotaped feedback? *The Journal of Business Communication*, Vol. 31, No.3, 213-219.
- Keppel, G. (1991). *Design and Analysis: A Researcher's Handbook*. Third Edition. Prentice Hall, Englewood Cliffs, New Jersey.
- Kihlstrom, J.F. & Cantor, N. (1984). Mental representations about the self. In L. Berkowitz (Ed.). *Advances in Experimental Social Psychology* (Vol.17, 1-47). New York: Academic Press.
- Kimble, C. and Helmreich, R. (1972). Self-esteem and the need for social approval. *Psychometric Science*, Vol. 26, No.6, 339-342.
- Kingstrom, P.O. & Mainstone, L. (1985). An Investigation of the Rater-Ratee Acquaintance and Rater Bias. *Academy of Management Journal*, Vol.28, No.3, 641-654.
- Kipnis, D., & Vanderveer, R. (1971). Ingratiation and the use of power. *Journal of Personality and Social Psychology*. Vol, 17, 280-286.
- Knight, F.B. (1923). The Effect of the "Acquaintance Factor" upon Personal Judgments. *Journal of Educational Psychology*, Vol.14, 129-142.
- Kolditz, T.A. & Arkin, R.M. (1982). An impression management interpretation of the self-handicapping strategy. *Journal of Personality and Social Psychology*, Vol.43, No.3, 492-502.
- Korman, A.K. (1966). Self-esteem variable in vocational choice. *Journal of Applied Psychology*, Vol.50, 479-486.

- Korman, A.K. (1970). Toward a hypothesis of work behavior. *Journal of Applied Psychology*, Vol.54, 31-41.
- Korman, A.K. (1976). Hypothesis of work behavior revisited and an extension. *Academy of Management Review*, Vol.1, No.1, 50-63.
- Kornhauser, A.W. (1926). What are rating scales good for? *Journal of Personnel Research*, Vol.5, 189-193.
- Kuiper, N.A. & Derry, P.A. (1982). Depressed and non-depressed content self-reference in mild depression. *Journal of Personality*, Vol.50, 67-79.
- Kunst-Wilson, W.R. & Zajonc, R.B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, Vol.207, 557-558.
- Larson, J.R.Jr., (1989). The dynamic interplay between employees' feedback-seeking strategies and supervisors' delivery of performance feedback. *Academy of Management Review*. Vol.14, 408-422.
- Latham, G.P. & Wexley, K.N. (1982). *Increasing Productivity through performance appraisal*, MA:Addison-Wesley.
- Landy, F.J. & Farr, J.L. (1983). *The measurement of work performance: Methods, theory, and applications*. New York, Academic press.
- Leary, M.R., & Kowalski, R.M. (1990). Impression Management: A literature review and two-component model. *Psychological Bulletin*, Vol. 107, No.1, 34-47.
- Levy, P.E., Albright, M.D., Cawley, B.D., & Williams, J.R. (1995). Situational and Individual determinants of feedback seeking: A closer look at the process. *Organizational Behavior and Human Decision Processes*, Vol.62, No.1, 23-37.
- Liden, R.C. & Mitchell, T.R. (1988). Ingratiation Behaviors in Organizational Settings. *Academy of Management Review*, Vol.13, No.4, 572-587.
- Locke, E.A. (1977). The Myths of behavior mod in organizations. *The Academy of Management Review*, Vol. 2, No. 4; 543-556.
- Lohdahl, T. & Kejner, M. (1965). The definition and measurement of job involvement. *Journal of Applied Psychology*, Vol.49, 24-33.
- London M., Wohlers, A.J. & Gallagher, P. (1990). 360-degree feedback surveys: A source of feedback to guide management development, *Journal of Management Development*, Vol.9, 17-31.

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

Longnecker, C.O., & Goff, S.J. (1990). Why performance appraisals still fail. *Journal of Compensation and Benefits*, Vol.6, No.3, 36-41.

Love, K.G. (1981). Comparison of Peer Assessment Methods: Reliability, Validity, Friendship Bias, and User Reaction. *Journal of Applied Psychology*, Vol.66, No.4, 451-458.

Mabe, P.A. & West, S.G. (1982). Validity of Self-Evaluation of Ability: A Review and Meta-Analysis. *Journal of Applied Psychology*, Vol.67, No.3, 280-297.

Martin, D.C. (1987). Factors influencing pay decisions: Balancing managerial vulnerabilities. *Human Relations*, Vol. 40, No.7, 417-430.

Maurer, T.J. & Tarulli, B.A. (1996). Acceptance of Peer/Upward Performance Appraisal systems: Role of work context factors and beliefs about managers' development capacity. *Human Resource Management*, Vol.35, No.2, 217-241.

Maurer, T.J., Raju, N.S. and Collins, W.C. (1998). Peer and Subordinate performance appraisal Measurement equivalence. *Journal of Applied Psychology*, Vol.83, No.5, 693-702.

Mayes, B.T.m & Allen, R.W. (1977). Toward a definition of organizational politics. *Academy of Management Review*, Vol.2, 672-677.

McCall, M.W., & Lombardo, M.M. (1983). *Off the track: Why and how successful executives get derailed*. (Tech. Report No. 21) Greensboro, NC: Center for creative leadership.

McFarlin, D.B., and Blascovich, J. (1981). Effects of self-esteem and performance on future affective preferences and cognitive expectations. *Journal of Personality and Social Psychology*, Vol.40, 521-531.

Miller, D.T. (1976). Ego involvement and attributions for success and failure. *Journal of Personality and Social Psychology*, Vol.34, 901-906.

Moreland, R.L. and Zajonc, R.B. (1979). Exposure effects may not depend on stimulus recognition. *Journal of Personality and Social Psychology*, Vol. 37, No.6, 1085-1089.

Morrison, E.W., & Bies, R.J. (1991). Impression management in the feedback-seeking process: A literature review and research agenda. *Academy of Management Review*, Vol.16, No.3, 522-541.

Moser, K., Schuler, H. & Funke, U. (1999) The moderating effects of raters' opportunities to observe ratees' job performance on the validity of an assessment center. *International Journal of Selection and Assessment*, Vol.7, 133-141.

Mossholder, K.W., Bedeian, A.G., & Armenakis, A.A. (1981). Role Perceptions, Satisfaction, and Performance: Moderating Effects of Self-Esteem and Organizational Level. *Organizational behavior and human performance*, Vol. 28, No. 2; 224-235.

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

Newcomb, T.M. (1959). Individual systems of orientation. In S. Koch (Ed.) *Psychology: A study of a science*, Vol.3. New York: McGraw-Hill, 384-422.

Northcraft, G.B., & Ashford, S.J. (1990). The preservation of self in everyday life: The effects of performance expectations and feedback context on feedback inquiry. *Organizational Behavior and Human Decision Processes*, Vol.47, 42-64.

Osgood, C.E. & Tannenbaum, P.H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, 42-55.

Payne, R.B. & Hauty, G.T. (1955). The effect of psychological feedback on work decrement. *Journal of Experimental Psychology*, Vol.50, 343-351.

Pierce, J.L., Gardner, D.G., Cummings, L.L. & Dunham, R.B. (1989). Organization-Based Self-Esteem: Construct Definition, Measurement, and Validation. *Academy of Management Journal*, Vol.32, No.3, 622-648.

Porter, L.W., Allen, R.W., & Angle, H.L. (1983). The politics of upward influence in organizations. In R.W. Allen & L.W. Porter (Eds.) *Organizational Influence Processes* (408-422). Glenview, IL: Scott, Foresman.

Roberts, G.E. (1998). Perspectives on enduring and emerging issues in performance appraisal. *Public Personnel Management*, Vol.27, No.3, 301-313.

Schlenker, B.R. (1980). *Impression Management*. Belmont, CA: Wadsworth.

Sederburg, M.E., & Rogelberg, S.G. (1998). Informed decisions: Research based practice notes- 360-degree feedback: Methodological advice from multiple sources. *The Industrial-Organizational Psychologist*, Vol.37, No.1, 67-76.



Sedikides, C. & Strube, M.J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. *Advances in Experimental Social Psychology*, Vol.29, 209-269.

Shavelson, R.J., Hubner, J.J., & Stanton, G.C. (1976). Self-concept: Validations of construct-interpretations. *Review of Educational Research*, Vol.46, 407-441.

Shrauger, J.S. (1975). Responses to evaluation as a function of initial self-perceptions. *Psychological Bulletin*, Vol.82, 581-596.

Shrauger, J.S. & Lund, A.K. (1975). Self-evaluation and reactions to evaluations from others. *Journal of Personality*, Vol. 43, No.1, 94-108.

Shrauger, J.S. & Rosenberg, S.E. (1970). Self-esteem and the effects of success and failure feedback on performance. *Journal of Personality*, Vol.38, 404-417.

Skowronski, J.J., Betz, A.L., Thompson, C.P., & Shannon, L. (1991). Social memory in everyday life: The recall of self-events and other events. *Journal of Personality and Social Psychology*, Vol. 60, 831-843.

Song, I.S., & Hattie, J. (1985). Relationships between self-concept and achievement. *Journal of Research in Personality*, Vol.19, 365-372.

Stanton, J.M. (1998). An empirical assessment of data collection using the internet. *Personnel Psychology*, Vol.51, 709-725.

Stepanski, K.M. & Fisicaro, S.A. (1999). The effects of method of rater selections on ratings in 360-degree feedback, *Paper presented at the 14<sup>th</sup> Annual Conference for the Society of Industrial and Organizational Psychology*, Atlanta, GA.

Strube, M.J. & Roemmele, L.A. (1985). Self-enhancement, self-assessment, and self-evaluative task choice. *Journal of Personality and Social Psychology*, Vol. 49, No.4, 981-993.

Summers, L. & Groehler, L. (2000). High Tech Adventures in boosting Response Rates. *Presentation at SIOP Conference*, April 2000.

Swann, W.B. & Read, S.J. (1981). Acquiring self-knowledge: The search for feedback that fits. *Journal of Personality and Social Psychology*, Vol. 41, No.6, 1119-1128.

Swann, W.B. Jr., Griffin, J.J. Jr., Predmore, S.C. & Gaines, B. (1987). The Cognitive-Affective crossfire: When self-consistency confronts self-enhancement. *Journal of Personality and Social Psychology*, Vol. 52, No.5, 881-889.

Swann, W.B., Jr. (1985). The self as architect of social reality. In B.Schlenker (Ed.), *The self and social life* (100-125). New York: McGraw-Hill.

Tabachnick, B.G. & Fidell, L.S. (1996). Using Multivariate Statistics, 3<sup>rd</sup> Edition. HarperCollins College Publishers.

Taylor, M.S., Fisher, C.D. & Ilgen, D.R. (1984). Individuals' reactions to performance feedback in organizations: A control theory perspective. In K.M. Rowland & G.R. Ferris (Eds.), *Research in Personnel and Human Resource Management*, Vol. 2, 81-124. Greenwich, CT: JAI Press.

Taylor, S.E. & Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, Vol. 103, 193-210.

Taylor, S.E. & Loebel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contracts. *Psychological Review*, Vol.4, 565-575.

Tedeschi, J.T. & Melburg, V. (1984). Impression management and influence in the organization. In S.B. Bacharach & E.J. Lawler (Eds.), *Research in the sociology of organizations*. Vol.2, 31-58. Greenwich, CT: JAI Press.

Tesser, A. (1986). Some effects of self-evaluation maintenance on cognition and action. In R.M. Sorrentino & E.T. Higgins (Eds.), *Handbook of motivation and cognition* (435-464). New York: Guilford Press.

Tharenou, P. (1979). Employees self-esteem: A review of the literature. *Journal of Vocational Behavior*, Vol.15, 1-29.

Tornow, W.W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management*, 32, 211-219.

Trope, Y. (1975) Seeking information about one's ability as a determinant of choice among tasks. *Journal of Personality and Social Psychology*. Vol 32, No.6, 1004-1013.

Trope, Y. (1979) Uncertainty-reducing properties of achievement tasks. *Journal of Personality and social psychology*, Vol 37, No.9, 1505-1518.

Trope, Y. (1986). Self-enhancement, self-assessment, and achievement behavior. In R.M. Sorrentino & E.T. Higgins (Eds.) *Handbook of motivation and cognition* (350-378). New York: Guilford.

Tsui, A.S. & Barry, B. (1986). Interpersonal Affect and Rating Errors. *Academy of Management Journal*, Vol. 29, No.3, 586-600

Van Velsor E., Leslie, J.B. and Fleenor, J.W. (1997) Choosing 360: A Guide to Evaluating Multi-Rater Feedback Instruments for Management Development. Greensboro, NC: Center for Creative Leadership

Wayne, S.J. & Ferris, G.R. (1990). Influence tactics, affect, and exchange quality in supervisor/subordinate interactions: A laboratory experiment and field study. *Journal of Applied Psychology*, Vol.75, No.5, 487-499.

Weiner, B. (1980). *Human Motivation*. New York: Holt, Rinehart, & Winston.

Weiss, H.M. (1977). Subordinate imitation of supervisory behavior. *Organizational Behavior and Human Performance*, Vol.19, 90-93.

Wells, L.E. & Marwell, G. (1976). *Self-esteem*. London: Sage Publications.

Westerman, J.W., & Rosse, J.G. (1997). Reducing the threat of rater nonparticipation in 360-degree feedback systems: An exploratory examination of the antecedents to participation in upward ratings. *Group & Organizational Management*, Vol. 22, No. 2, 288-309.

Wike, E.L., & Church, J.D. (1982). Nonrobustness in F tests:1. A replication and extension of Bradley's study. *Bulletin of the Psychometric Society*, Vol.20, 165-167.

Wylie, R.C. (1974). *The self-concept: A review of methodological considerations and measuring instruments*. Lincoln: University of Nebraska Press.

Yammarino, F.J., & Atwater, L.E. (1993). Understanding self-perception accuracy: Implications for human resources management. *Human Resource Management*, Vol, 32, 231-247.

Young, S.A., Daum, D.L., Robie, C., & Macey, W.H. (2000). Paper versus Web Survey Administration: Do different methods yield different results? *Paper presented at the 15<sup>th</sup> annual conference of the Society for Industrial and Organizational Psychology*, New Orleans.

Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, Vol.35, No.2, 151-175.

## **APPENDIX 1**

To whom this may concern,

I am a graduate student at Concordia University pursuing my Masters in Administration, (management option) and would like your help in collecting data for my thesis. My thesis focuses on the different strategies that individuals use in a multi-source feedback process. Multi-source feedback is an increasingly popular assessment method and it is sometimes referred to as 360-degree feedback. When individuals participate in a multi-source feedback process at work, they are often given the opportunity to choose who among their many peers and subordinates will rate their behavior. Given this, the potential for bias and the self-selection of results is very probable.

Therefore, in order to further investigate this issue, we have developed a web-based survey that takes approximately 20-30 minutes to complete. We are seeking individuals who work, or have worked **full-time** in the last 2 years and have/had a minimum of **4 subordinates** under their supervision to complete the survey. Since business students often have both an academic and managerial orientation, we have targeted our sample selection towards MBA and EMBA students. This is where you come in, I would greatly appreciate it if you could either provide me with the e-mail addresses of your MBA and EMBA students or forward them a message that requests their participation.

The questions in the survey are very straightforward and participants will not be asked to identify themselves at any time during the survey, therefore all data will remain anonymous. The project has already been approved by Concordia's human research ethics review committee. I would be glad to send you a copy of my thesis proposal for your review. Also, you can view the survey by clicking on the link below:

<http://mana-research.concordia.ca/brutus/survey.html>

You can let me know if you are interested in participating by either calling me at the telephone number listed above or by e-mailing me. At that point I can answer and additional questions you may have.

Thank you for your time and consideration,

Sandra Petosa (primary investigator), under the supervision of Stéphane Brutus, Ph.D.

## **APPENDIX 2**

To whom this may concern,

I am a graduate student at Concordia University pursuing my Masters in Administration, (management option) and would like your help in collecting data for my thesis. My thesis focuses on the different strategies that individuals use in a 360-degree feedback process and in order to investigate this further, we have developed a web-based survey that takes approximately 20 minutes to complete.

Given this, I am imploring MBA and EMBA students to participate. I am seeking individuals who work, or have worked **full-time** in the last 2 years and have/had a minimum of **4 subordinates** under their supervision. I would greatly appreciate it if you could take the time to complete the survey. The questions in the survey are very straightforward and you will not be asked to identify yourself at any time during the survey, therefore all data will remain anonymous. Once the study is complete a report which details the findings of the study can be made available to all interested participants.

You can complete the survey at your convenience by clicking on the link below:

<http://mana-research.concordia.ca/brutus/survey.html>

Thank-you for you time and consideration,

Sandra Petosa (primary investigator), under the supervision of Stéphane Brutus,  
Ph.D.

### APPENDIX 3



#### Welcome to our online Survey!

This is a survey developed by The Faculty of Commerce and Administration at Concordia University.

Completion of the survey is voluntary but will provide us with a better understanding of the strategies that individuals use in a multi-source feedback process. The survey will take 15-20 minutes to complete.

There are no right or wrong answers so please be completely honest if you choose to participate.

SUBMIT AND CONTINUE

### **Consent form to participate in research**

This is to state that I agree to participate in the program of research being conducted by Sandra Petosa in the M.Sc Administration department of Concordia University.

#### **A. Purpose**

The purpose of the research is to determine the strategies that individuals use in a multi-source feedback process.

#### **B. Procedures**

You will complete the questionnaire that follows (it will take approximately 15-20 minutes to complete). As you move through the pages of the survey, you will be asked to submit your data to our database. You will not be asked to identify yourself at any time during the survey, therefore your anonymity is ensured.

#### **C. Conditions of participation**

- I understand that I am free to withdraw my consent and discontinue my participation at any time without negative consequences.
- I understand that my participation in this study is anonymous.
- I understand that the results of this study may be published.
- I understand the purpose of this study and know that there is no hidden motive of which I have not been informed.

I HAVE CAREFULLY READ THE ABOVE AND UNDERSTAND THIS AGREEMENT. I FREELY CONSENT AND VOLUNTARILY AGREE TO PARTICIPATE IN THIS STUDY

I AGREE      I DISAGREE

**1**The questions in this survey will ask you to refer to a time when you were working full-time and had subordinates under your supervision. Since it is VERY IMPORTANT THAT THE SAME JOB/POSITION BE REFERRED TO THROUGHOUT THIS SURVEY, please indicate the position you held and the name of the company that you worked for.

Position held:

Company name:

*(if you do not want to give the company name you can use a nick-name or acronym instead)*

**2**How long ago did you hold that position?

- ☒ Current
- ☐ 1-3 months ago
- ☐ 4-6 months ago
- ☐ 7-9 months ago
- ☐ 10-12 months ago
- ☐ 12+ months ago

SUBMIT AND CONTINUE



PART 1

**1**Sex: ☐ Male ☐ Female

**2**Age:

**3**Number of years in school:   
(beginning with elementary school)

**4**Highest degree earned (Mark one)

- ☐ High School
- ☐ College
- ☐ Bachelor's
- ☐ Master's
- ☐ Doctorate/Professional

**5**Type of organization you work for  
(Mark only one under ONE appropriate sector)

A. Business Sector

- ☐ Manufacturing
- ☐ Transportation, Communication, Utilities
- ☐ Wholesale/Retail Trade
- ☐ Finance, Insurance, Banking
- ☐ Health

Other:

**B. Public Sector**

- ☐ Elementary and Secondary Education
- ☐ Higher Education
- ☐ Military
- ☐ Government Agency

Other: \_\_\_\_\_

**C. Private Nonprofit Sector**

- ☐ Elementary and Secondary Education
- ☐ Higher Education
- ☐ Health
- ☐ Human Services

Other: \_\_\_\_\_

**6 Organizational Level at XYZ (Mark One)**

- ☐ TOP-Chief Executives, Operating Officers, Presidents
- ☐ EXECUTIVE-Vice Presidents, Directors, Board-level Professionals
- ☐ UPPER MIDDLE-Department Executives, Plant Managers, Senior Professional Staff
- ☐ MIDDLE-Office Managers, Professional Staff, Middle-Level Administrators
- ☐ FIRST LEVEL-Forepersons, Crew Chiefs, Section Supervisors
- ☐ HOURLY EMPLOYEES-Machine Operators, Clerical/Secretarial and Support Staff, Technicians
- ☐ NOT RELEVANT IN MY SITUATION

**7 Your Function at XYZ**

*(Mark the one that is most closely related to your work)*

- ☐ Accounting
- ☐ Administration
- ☐ Advertising/ Public Relations
- ☐ Credit/ Finance
- ☐ Education
- ☐ Engineering
- ☐ HR/ Training
- ☐ Information Systems/ Data Processing
- ☐ Law
- ☐ Manufacturing

- ☐ Marketing
- ☐ Materials Management/ Purchasing
- ☐ Medicine
- ☐ Operations
- ☐ Product Development
- ☐ Quality Control
- ☐ Research/ Analysis
- ☐ Research and Development
- ☐ Sales
- ☐ Secretary/ Support
- ☐ Security
- ☐ Social Service
- ☐ Systems Analysis
- ☐ Top Management

Other:

SUBMIT AND CONTINUE

## PART 2

In reference to your job at XYZ, indicate the extent to which you agree/disagree with the following statements:

	<b>Strongly Agree</b>	<b>Agree</b>	<b>Neither Agree nor Disagree</b>	<b>Disagree</b>	<b>Strongly Disagree</b>
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>1</b> I count at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>2</b> I am taken seriously at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>3</b> I am important at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>4</b> I am trusted at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>5</b> There is faith in me at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>6</b> I make a difference at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>7</b> I am valuable at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>8</b> I am helpful at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>9</b> I am efficient at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>10</b> I am cooperative at XYZ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SUBMIT AND CONTINUE

### PART 3

Now, name 4 co-workers (same hierarchical level) that you worked with at XYZ:

**1**Name :

**2**Name :

**3**Name :

**4**Name :

SUBMIT AND CONTINUE

**1**How long have you worked with Diana? \_\_\_\_\_ years

**2**How much do you like Diana?

I don't like this person at all	I don't like this person	I neither like nor dislike this person	I like this person	I like this person very much
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**3**How frequently do you have contact with Diana?

Not Frequently at all	Not Frequently	Neutral	Frequently	Very Frequently
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**4**If Diana were to rate your performance, what kind of evaluation would he/she give you overall?

Highly Negative	Negative	Neutral	Positive	Highly Positive
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**5**I get along with Diana.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**6**How familiar is Diana with your day-to-day tasks?

Not Very Familiar	Unfamiliar	Neutral	Familiar	Very Familiar
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**7** Working with Diana is a pleasure.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**8** I consider Diana a good friend.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

SUBMIT AND CONTINUE

(this page is repeated for each of the 3 remaining co-workers)

#### PART 4

Now, name 4 subordinates that you supervised (lower hierarchical level) at XYZ:

**1**Name :

**2**Name :

**3**Name :

**4**Name :

SUBMIT AND CONTINUE



**1**How long have you worked with Domenica? \_\_\_\_\_ years

**2**How much do you like Domenica?

I don't like this person at all	I don't like this person	I neither like nor dislike this person	I like this person	I like this person very much
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**3**How frequently do you have contact with Domenica?

Not Frequently at all	Not Frequently	Neutral	Frequently	Very Frequently
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**4**If Domenica were to rate your performance, what kind of evaluation would he/she give you overall?

Highly Negative	Negative	Neutral	Positive	Highly Positive
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**5**I get along with Domenica.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**6**How familiar is Domenica with your day-to-day tasks?

Not Very Familiar	Unfamiliar	Neutral	Familiar	Very Familiar
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**7** Working with Domenica is a pleasure.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

**8** I consider Domenica a good friend.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5
⌋	⌋	⌋	⌋	⌋

SUBMIT AND CONTINUE

(this page is repeated for each of the 3 remaining subordinates)

PART 5

Which 2 of the 4 **CO-WORKERS** you described would you choose in a performance evaluation if the ratings were going to be **ANONYMOUS AND PRIVATE** (i.e., no one but yourself, not even your immediate supervisor would see the results)?

Diana  
☐

Walter  
☐

Lina  
☐

Marco  
☐

Which 2 of the 4 **CO-WORKERS** you described would you choose in a performance evaluation if the ratings were going to be public (i.e., the results would be available to your supervisor and could influence decisions related to **COMPENSATION, PROMOTION, TERMINATION OR LATERAL TRANSFERS**)?

Diana  
☐

Walter  
☐

Lina  
☐

Marco  
☐

---

SUBMIT AND CONTINUE

PART 6

Which 2 of the 4 **SUBORDINATES** you described would you choose in a performance evaluation if the ratings were going to be **ANONYMOUS AND PRIVATE** (i.e., no one but yourself, not even your immediate supervisor would see the results)?

Domenica  
☐

Jess  
☐

Laura  
☐

Adam  
☐

Which 2 of the 4 **SUBORDINATES** you described would you choose in a performance evaluation if the ratings were going to be public (i.e., the results would be available to your supervisor and could influence decisions related to **COMPENSATION, PROMOTION, TERMINATION OR LATERAL TRANSFERS**)?

Domenica  
☐

Jess  
☐

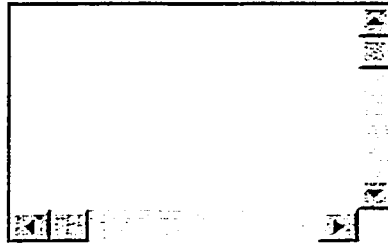
Laura  
☐

Adam  
☐

SUBMIT AND CONTINUE

You have completed the survey!

In the box below you can provide us with your comments about the strategy that guided your decision in choosing raters or with any other general comments you may have. If you would like to be entered in the draw for \$150 please send you name and telephone number to Sandra Petosa



SUBMIT

CLEAR