Brigham Young University
BYU ScholarsArchive

Theses and Dissertations

2009-02-19

Predicting Transformational Leadership: Self-Other Agreement in Multi-Source Feedback

Ryan Hamilton Shatzer
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

 Part of the Psychology Commons

BYU ScholarsArchive Citation

Shatzer, Ryan Hamilton, "Predicting Transformational Leadership: Self-Other Agreement in Multi-Source Feedback" (2009). *Theses and Dissertations*. 2010.
https://scholarsarchive.byu.edu/etd/2010

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

PREDICTING TRANSFORMATIONAL LEADERSHIP:

SELF-OTHER AGREEMENT IN MULTI-SOURCE FEEDBACK

by

Ryan H. Shatzer

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Psychology

Brigham Young University

April 2009

Copyright © 2009 Ryan H. Shatzer

All Rights Reserved

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Ryan H. Shatzer

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Bruce L. Brown, Chair
Date	David J. Cherrington
Date	Ross A. Flom

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Ryan H. Shatzer in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date	Bruce L. Brown Chair, Graduate Committee

Accepted for the Department

Date	Harold L. Miller, Jr. Graduate Coordinator

Accepted for the College

Date	Susan Rugh Associate Dean, College of Family, Home and Social Sciences

ABSTRACT

PREDICTING TRANSFORMATIONAL LEADERSHIP:

SELF-OTHER AGREEMENT IN MULTI-SOURCE FEEDBACK

Ryan H. Shatzer

Department of Psychology

Master of Science

Multi-Source Feedback (MSF) has become an important tool for leadership development programs. Previous research has examined how self-other agreement in MSF relates to leadership effectiveness. Discrepancies exist in the literature between how to measure self-other agreement and which method best depicts self-awareness. The current study examined the relationship between various measurements of self-other agreement, self-awareness and transformational leadership. MSF data were collected from target leaders (n = 31), and their respective direct report, peer and supervisor raters (n = 233). Raters also evaluated their leaders' self-awareness and leadership behavior. Self-other agreement was measured using a reliability coefficient, *self-other agreement r*, and a difference squared score, *self-other agreement* D^2. These measures of self-other agreement as well as the direct measure of self-other awareness

were used to predict transformational leadership. Results indicated that *self-other agreement r* did not significantly predict transformational leadership, while *self-other agreement* D^2 did significantly predict some of the dimensions of transformational leadership. However, the direct measure of self-awareness was the strongest predictor of transformational leadership. The two methods of calculating self-other agreement did not have a significant correlation, indicating that they may be measuring different constructs. The direct measure of self-awareness also did not correlate significantly with self-other agreement, suggesting that there is a conceptual gap between these two constructs and complexities may arise when researchers operationalize self-other agreement as self-awareness. The issues surrounding the various methods of measuring self-other agreement, as well as the possible confounding effects of the direct measure of self-awareness and difference scores are discussed. Implications for interpreting self-other agreement in MSF processes are also discussed.

ACKNOWLEDGEMENTS

Special thanks to the members of my thesis committee that have provided valuable insight, contributions and feedback: Drs. Bruce Brown, Ross Flom and David Cherrington. Dr. Brown, you have been very supportive and flexible as I have explored alternative fields within psychology. Thanks also to my cohorts in the psychology department, and those at the business school that have so kindly accepted me. I would also like to acknowledge the HR managers of the four companies that assisted me in this project. Thank you.

Many have contributed to this accomplishment outside the circles of academia. Much of what I have accomplished has been through the support of my parents. Thank you for what you have taught me and helped me to become. Unending thanks as well to my siblings, in-laws and grandparents that have helped and supported me along the way. As always, thank you Meredith for your love and support through this whole experience. Thank you for your patience as I spent many evenings away in the grad lab or on trips throughout the valley to do *360s*. I hope that I can be worthy of your love and faith in me. I dedicate this to you and to our baby: I will do my best to preside in love and righteousness, the essence of what leadership really is.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

In corporations worldwide the demand for leadership talent far exceeds the supply. In a study by Mckinsey & Co. (Chambers, Foulon, Hansfield-Jones, Hankin, & Micheals, 1998), 40 percent of the corporations studied claimed they cannot pursue the organization's growth strategy because of the lack of leadership talent. Many top consulting firms have recommended a "grow your own" approach of retaining and developing leaders within the organization (Gretchko, 2007). A core aspect of this internal leadership development process includes feedback evaluations. One popular method is multi-source feedback (MSF), also known as 360-degree feedback, where several raters from various levels of the organization give feedback to a target leader. The current literature within industrial/organizational psychology has attempted to find the most efficient ways to develop, interpret, and implement MSF in corporate practice to develop leaders.

The use of MSF for leadership development and performance evaluation has become widespread (Atwater & Brett, 2006). MSF generally includes feedback from supervisor, peer, self and direct report raters (Foster & Law, 2006). Multiple sources of raters take into account a wider range of employee behavior for a more complete picture of feedback needs (Borman, 1997). The MSF process usually consists of multiple sources rating performance competencies (e.g. productivity, communication) of a target leader in an organization. The ratings from each source are presented in a report to the target leader and typically used for leadership development purposes. Ideally, MSF should help employee's see themselves as others see them, and provide them with developmental feedback about needed behavioral changes (Atwater & Waldman, 1998).

The MSF process also allows for comparing a leader's self-ratings with the ratings of others. Researchers have used the level of agreement between leaders' self-ratings and other's ratings to predict individual outcomes (Atkins & Wood, 2002; Atwater & Waldman, 1998). Self-other agreement is the degree to which an individual's perception of performance aligns with other's perceptions of that individual's performance. The relationship between self-other agreement has been found to be quite complex, with several factors contributing to other's ratings and self-ratings. Elements of other's ratings, self-ratings, and self-other agreement will be discussed.

Other's Ratings

Other's ratings include supervisor, peer, and direct report competency ratings of the target leader. The ratings are a measure of other's perceptions and are subject to perception bias. Because of this bias, other's ratings should not necessarily be seen as true scores (Alimo-Metcalfe, 1998). Several factors need to be taken into account when interpreting the ratings of others in MSF. In general the greater the opportunity to observe the leader's performance, the greater will be the accuracy of the ratings (Rothstein, 1990; Sundvik & Lindeman, 1998). Conway (1996) has shown that rater's perception bias decreases as opportunities to observe leader's behavior increase. Also, Rothstein (1990) has found that agreement between supervisor's ratings increase with ratee tenure in an organization. Warr and Bourne (1999) found that correlations between self-other ratings were higher for behaviors that were rated as more observable by supervisors.

The acquaintanceship between the leader and the raters can also affect other's ratings. Other's ratings more closely reflect assessment center evaluations once the leader and rater have worked together for approximately two years (Moser, Schuler, & Funke, 1999). Agreement between peer's ratings increased when the rater and ratee were more acquainted, suggesting

acquaintance leads to more accurate ratings among peers (Mumford, 1983). However, length of acquaintance with a supervisor can also lead to an overestimation of ratee performance (Sundvik & Lindeman, 1998), and leniency effects (Furnham & Stringfield, 1998). There is a point of diminishing returns in acquaintanceship length, and an actual decrease in accuracy with extensive acquaintanceship (Sundvik & Lindeman, 1998). It has been suggested that the rater should be familiar enough with the target leader to know their ability, but not so acquainted that friendships interfere with accurate ratings due to leniency effects (Moser, et al., 1999).

Raters who are higher performers tend to evaluate more strictly than low performers (Saavedra & Kwun, 1993). Saavedra and Kwun have suggested that this difference in rating can be attributed to social comparisons. Raters tend to compare the performance of others to their own performance, and when the rater's performance is high there may also be a higher standard for evaluation. Situational factors such as job factors, previous rating experiences, and similarities to the leader also influence other's ratings (Yammarino & Atwater, 1997). Also, certain job functions naturally receive more feedback than other positions (e.g. the on stage performer as opposed to the backstage technician). Individual characteristics of the rater, such as analytic ability, memory, interpersonal orientation, and self-esteem, also influence their ratings (Yammarino & Atwater, 1997).

Supervisor, peer, and direct report ratings are based on complex factors and are subject to each rater's perception biases. These factors need to be considered when selecting raters and calculating self-other agreement. Rater biases may also contribute to greater self-other discrepancies. Two controls in MSF can be used to reduce perception bias. The aggregate of other's scores may control for individual factors and biases. Also, multiple raters from multiple levels of the organization will encompass more of the ratee's behavior and help reduce bias.

Based on this reasoning, supervisor, peer, and direct report ratings are aggregated to create a mean "other's" score.

Self-Ratings

Self-ratings are the target leader's ratings of his or her own performance, using the same competencies and questions as the other raters. Self-ratings can be thought of as a form of self-perception (Van Velsor, Taylor, & Leslie, 1993). There are several biases in self-ratings, the most prominent is self-enhancement. In general people consistently tend to over rate their performance (Atwater & Yammarino, 1992; Harris & Schaubroeck, 1988). Leaders that have worked longer in their organization are more likely to inflate their self-ratings, and are less likely to be receptive to feedback (Sinha, 2004). Newer leader's self-ratings tend to agree more with their supervisor ratings than tenured leader's self-ratings (Sinha, 2004). Higher level management (e.g. senior executives) tend to have greater self-other discrepancies than lower level management (Sala, 2003). Sala suggested that the larger discrepancies in upper management and more tenured leaders may be due to their assumption that experience and position are equivalent to performance. These higher discrepancies may also be due to the organization's culture, where subordinates are reluctant to provide corrective feedback to high level executives (Sala, 2003).

The amount of previous experience or practice in assessing personal performance may affect the accuracy of self-ratings. Previous experience with self-assessments heightens self-awareness and self-other accuracy (Sinha, 2004). Those who receive more feedback on the job are less likely to have self-other discrepancies (Atwater & Brett, 2006). Stress and job pressures associated with the performance ratings can contribute to self-accuracy (Yammarino & Atwater, 1997). Situational control, the degree to which managers feel they are confident and in control of what they are doing, affects self-other discrepancies. A study collecting MSF ratings and

measurements of situational control of 92 managers found that when situation control was high, self-other discrepancies were lower (Baril, Ayman, & Palmiter, 1994).

Personality and gender may also affect self-other accuracy. Academically proficient individuals are more likely to accurately self-rate their own performance in academic settings (Yammarino & Atwater, 1997). Individuals higher in intelligence and short-term memory also tend to give more accurate self-ratings (Yammarino & Atwater, 1997). In one study, individuals higher in intelligence, achievement status and internal locus of control were associated with more agreement in self-evaluations (Mabe & West, 1982). Women have a tendency to rate their performance lower than males, and consequently have less self-other discrepancies (Alimo-Metcalfe, 1998; Atwater & Waldman, 1998; Atwater & Yammarino, 1993).

Self-ratings tend to have low correlations with other sources of ratings. Harris and Schaubroeck's (1988) meta-analysis of 54 studies showed that self-other correlations averaged .35, while other-other correlations were .62. From a sample of 1460 managers, the correlations among other raters averaged .40, while self-other correlations averaged .25 (Atwater & Waldman, 1998). When comparing self-rating with direct report ratings, London and Wholers (1991) found a correlation of .06, concluding marginal agreement between self and direct report ratings. In a more recent study, the average correlations between self and supervisor, peers, and direct report's ratings were .19, while the average correlations between other's ratings were .30 (Sala & Dwight, 2002).

Because of the tendency for individuals to over rate their own performance, self-ratings are poor predictors of outcome measures. Self-ratings, when compared to supervisor, peer and direct report ratings, have the lowest correlation with the performance measures of production count, sales, promotion, and managerial performance appraisals (Atwater & Waldman, 1998;

Harris & Schaubroeck, 1988; Yammarino & Atwater, 1997). Maybe and West (1982), in a meta-analysis, found there was an average of .04 correlation between manager's self-rating and annual performance appraisals. These studies conclude that self-ratings and ratings from other are different constructs.

Since self-ratings do not predict how well a manager will perform, other's ratings are seen as a more useful measure. Some researchers have concluded that self-ratings have little beneficial value (Fleenor, McCauley, & Brutus, 1996). Other researchers argue that self-rating may be useful when compared to other's ratings as a measure of self-other agreement.

Self-Other Agreement

Self-other agreement is the degree to which an individual's perception of performance aligns with other's perceptions of their performance. Researchers have found that the degree and type of self-other agreement is relevant to outcome measures. Four levels of self-other agreement in the literature have emerged: over-estimators, in-agreement/good, in-agreement/poor, and under-estimators.

Over-Estimators. These are individuals that over rate their performance on MSF feedback when compared to other's ratings. Theoretically, over-estimators tend to over rate their performance because they are unaware of how they are perceived by others (Yammarino & Atwater, 1997). Over-estimators are characteristic of setting unrealistically high goals for employees and themselves (Fleenor, et al., 1996). Career derailment is a problem for over-estimators because they are less open to feedback and they feel their performance level is already high (Atwater & Yammarino, 1993; Fleenor, et al., 1996). In terms of leadership, over-estimators are likely to have poor relationships with direct reports, and are less likely to receive recommendations for promotion (Atwater & Waldman, 1998; Atwater & Yammarino, 1993). Over-estimators received the lowest ratings regarding their managerial practices in comparison

to in-agreement and under-estimators (Van Velsor, et al., 1993). Over-estimators tend to be egotistical and arrogant, feel underappreciated and under recognized, rationalize negative feedback, and accept positive feedback as more accurate (Atwater & Waldman, 1998; Yammarino & Atwater, 1997). Those with large self-other discrepancies seem to misjudge their self-efficacy. Bandura (1982) has stated that:

> Acting on misjudgments of personal efficacy can produce adverse consequences, accurate appraisal of one's own capabilities has considerable functional value. Self-efficacy judgments, whether accurate or faulty, influence choice of activities and environmental settings. People avoid activities that they believe exceed their … capabilities, but they undertake and perform assuredly those that they judge themselves capable of managing. (p. 123)

This statement suggests that over-estimators will attempt to perform tasks that they are ill-prepared to do, take unnecessary risks, and overlook more realistic task options. Over-estimators lead to diminished individual and organizational outcomes (Yammarino & Atwater, 1993).

In-Agreement/Good. In-agreement/good raters are defined as those who do not differ significantly with other's ratings, and who receive above average ratings from others (Yammarino & Atwater, 1997). In-agreement/good raters are more likely to obtain higher positions and are more recognized by supervisors for promotions (Bass & Yammarino, 1991). These type of self-raters also tend to receive feedback positively, be successful managers and leaders, and have fewer conflicts with others (Yammarino & Atwater, 1997). These self-aware managers are better able to associate with employees, and therefore, install trust and commitment. In a study measuring leadership behaviors, those that were in agreement with other's ratings were found to have higher levels of trust from the organization (Sosik, 2001). In-

agreement/good individuals are ideal employees who enhance organizational outcomes (Yammarino & Atwater, 1993).

In-Agreement/Poor. In-agreement/poor raters are individuals whose self-ratings are in agreement with other's ratings, but both ratings are below average (Yammarino & Atwater, 1997). These raters recognize their weaknesses, and acknowledge other's awareness of those weaknesses. These raters are characterized by having lower productivity, knowledge, skills, and abilities for the job (Yammarino & Atwater, 1997). Smircich and Chesser (1981), suggest that in-agreement/poor individuals are still preferable to over-estimators because it indicates that they have some awareness of their poorer performance. Feedback for these individuals often leads to negative outcomes, with few intentions to improve performance (Smircich & Chesser, 1981).

Under-Estimators. These individuals are characterized as self-rating significantly lower than other's ratings (Yammarino & Atwater, 1997). These raters are either overly modest or do not realize their strengths. They have mixed performance abilities, tend to have lower aspiration levels, and display lower self-esteem (Yammarino & Atwater, 1997). Researchers suggest under-estimators may be hard workers (to compensate for perceived poor performance), pass up options for which they would be qualified for, and avoid risk taking (Atwater & Waldman, 1998). Under-estimators seem to be mixed in their performance and abilities. For example, promotion was positively correlated to in-agreement raters, negatively correlated to over-estimators, and unrelated to under-estimators (Atwater & Waldman, 1998). Results of other studies investigating under-estimators are either non-significant or contradictory, showing mixed results (Atwater & Waldman, 1998; Sosik, 2001).

Measuring Self-Other Agreement

Methodological issues in calculating self-other agreement have complicated this current line of research. The categorization approach, which was the primary method of calculating self-

other agreement until recent years, has been criticized as being confounding (Atkins & Wood, 2002). When other's ratings are used to measure leader's performance and used to define post hoc categorization into agreement groups, it is only logical that under-raters perform better than over-raters. By definition the over-rater group would not include anyone that received high scores from others. Given the narrow variation in self-rating and ceiling effect (James, Demaree, & Wolf, 1993), in order to over-estimate one's performance one must receive low performance ratings. Conversely, under-raters would not receive low scores from others. The current categorization method seems to be a measure of performance more than a measure of self-other agreement (Fleenor, et al., 1996). Other methods that get at the pattern of self-other agreement may better capture the construct of self-other agreement.

Polynomial regression, recommended by Edwards (1995), attempts to better capture the construct of self-other agreement. Polynomial regression allows one to see the form of agreement between self-ratings and other's ratings compared to another outcome measure. Difference score approach, on the other hand, constrains the functional form of the relationship between self-other scores (Atkins & Wood, 2002). Polynomial regression treats self-ratings and other ratings as two separate constructs, rather than turning the self-other difference into a univariate model as with difference scores. Results, which produce a three-dimensional graph giving a rich display of information, shows the numeric value of the dependant variable along the line of perfect self-other agreement.

Polynomial regression is the correct method for measuring stereotype accuracy; however, differential accuracy demonstrates a more meaningful self-other agreement index (Cline & Richards, 1960). Stereotype accuracy is the ability to rate performance generally, while differential accuracy is the ability to accurately determine an individual's strengths and

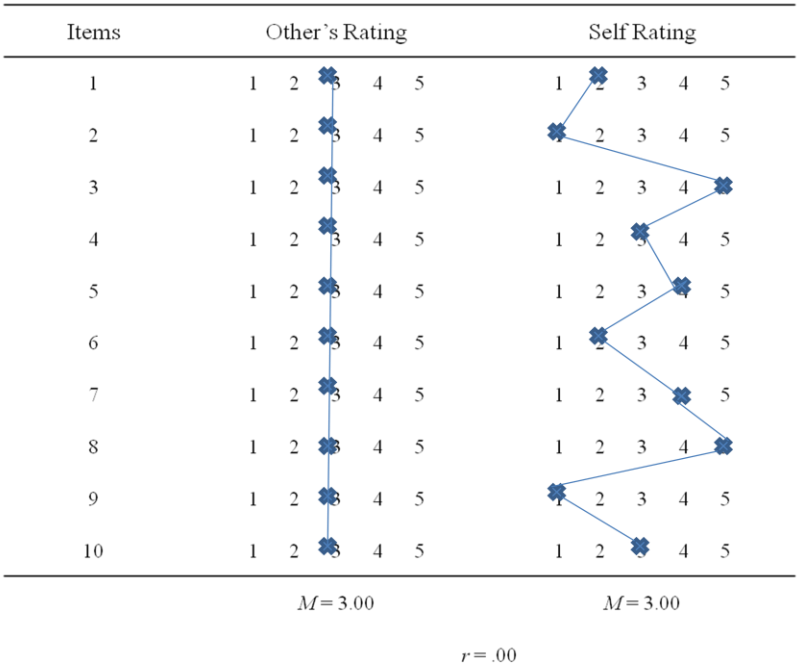

Figure 1. Data representing perfect self-other agreement with stereotype accuracy.

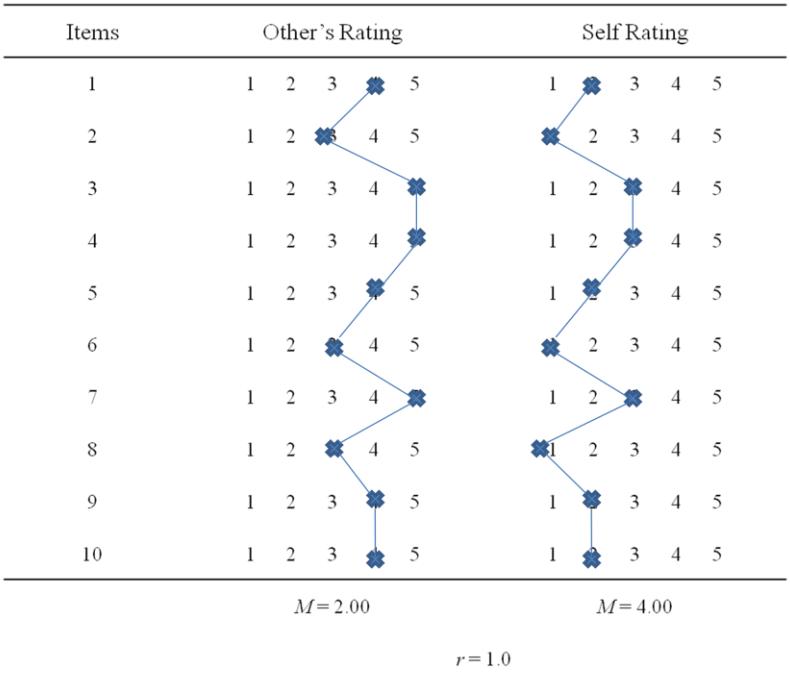

Figure 2. Data representing perfect self-other agreement with differential accuracy.

weaknesses comparatively (Bernieri, Zuckerman, Koestner, & Rosenthal, 1994; Cline & Richards, 1960). When computing the self-other agreement using the polynomial regression technique a total aggregate is calculated for the self and other scores. Total aggregate scores do not capture the pattern of agreement between a self-rater and other raters on the various items. For example, in Figure 1 the self-ratings do not deviate, while the other's scores deviate substantially. Although self-ratings and other's ratings did not seem to agree in regards to the pattern of accuracy (differential accuracy), the total aggregate (stereotype accuracy) showed perfect agreement.

Differential agreement better captures the pattern of agreement, taking into account how other-ratings and self-ratings vary together. This technique uses the average covariance divided by the average variance to form a reliability coefficient, similar to Intra-Class Correlation Coefficients (ICCC). Figure 2 gives an example of perfect differential accuracy agreement, where the self-ratings vary on the 5-point scale similar to other's ratings.

Self-Awareness

Four competing theories have been proposed to explain the self-other discrepancy, which are egocentric bias theory, self-enhancement theory, self-presentation theory, and self-awareness theory. Although there is considerable overlap between these theories, each will be considered separately. Egocentric bias explains the self-other discrepancy as rater's tendency to inflate their self-ratings (Harris & Schaubroeck, 1988). This inflation of self-ratings is caused by the fundamental attribution error. Individuals tend to attribute their good performance to dispositional factors and attribute poor performance to situational factors (Farh & Dobbins, 1989). While from the other perspective, others attribute good performance to situational factors and poor performance to dispositional factors (Harris & Schaubroeck, 1988). It is also

suggested that direct reports, peers and supervisors have similar cognitive evaluation techniques due to similar frames of reference (Farh & Dobbins, 1989).

The self-enhancement theory suggests that people seek and prefer favorable feedback. According to this theory, people are motivated to enhance their self-esteem by seeking favorable feedback in order to achieve a high level of personal worth (Dauenheimer, Stahlberg, & Petersen, 1999). Feedback that is favorable leads to positive affective states, unfavorable feedback leads to negative affective states (Katz, Arias, & Beach, 2000). Two versions of the self-enhancement theory have emerged in the literature. Simple self-enhancement is a universal motive in people to promote the perception that others think well of them (Swann, Griffin, Predmore, & Gaines, 1987). Compensatory or defensive self-enhancement refers to the assumption that people with negative self-concepts rarely receive positive feedback and then compensate for that discrepancy by promoting themselves (Swann, et al., 1987). The distinct difference between these two versions is that the former assumes that all people are motivated to self-enhance, whereas the later assumes that people with negative self-concepts are more motivated to self-enhance than people with a positive self-image (Swann, Pelham, & Krull, 1989).

It is somewhat healthy to have a small degree of self-enhancement bias for adjustment, but if this upward biases causes one to ignore weaknesses and failures it will negatively impact performance (Yammarino & Atwater, 1997). The application of this theory would mean that individuals with self-other discrepancies downplay their weaknesses and over estimate their strengths to feel better about themselves. The danger in this viewpoint is that a false sense of self-accomplishment can lead people to accept difficult tasks in fields which they are not

competent, as well as overlook risk factors and weaknesses in situations where failure is probable (Atwater & Waldman, 1998; Bandura, 1982).

Different from self-enhancement theory, self-presentation theory states that people over rate self-performance to present themselves in a positive light to others (Sinha, 2004). These individuals are high self-monitors and attempt to control the perceptions others form of them (Snyder, 1974). This is accomplished by controlling information in social settings by monitoring negative behavior and promoting their positive attributes. The application of this theory suggests that individuals may over-estimate in order to present a positive view of themselves to others, accentuating positive attributes while masking negative attributes.

Self-awareness theory recognizes self-enhancement as part causal, but primarily attributes the self-other discrepancy to a misunderstanding of other's perceptions. Individuals with self-other discrepancies are ignorant of other's perception of their performance, and have low self-awareness. Self-awareness theory (Duval & Wicklund, 1972) rests on the basis that one's focus is dichotomous. The two states are what Duval and Wicklund call objective self-awareness (an internal focus of attention on the self), and subjective self-awareness (an external focus on environmental factors). The term objective was used to describe one's internal attention because they are the object of their consciousness. Subjective denotes one as the subject of their consciousness, when attention is directed to external factors.

According to the self-awareness theory, when individuals are objectively self-aware their conscious attention is primarily focused on the self, or the aspect of the self that is most salient for the time being. During this time of self-focus, individuals self-evaluate their thoughts, emotions, and behaviors. These individuals are more aware of their inner self than individuals that have an external focus. Conversely, objective self-aware individuals are less aware or

concerned with external environmental factors. Objective self-aware individuals have a more accurate view of their behavior, and can therefore give a more accurate self-report (Gibbons, 1983). Current research in industrial/organizational psychology focuses primarily on objective self-awareness and refers to this internal focus of attention as simply self-awareness.

Self-aware individuals are able to self-observe, and accurately compare their behavior to a standard or social norm (Atwater & Yammarino, 1992). Self-awareness also encompasses the ability to be aware of other's perceptions (receiving feedback from others) and incorporate other's perceptions into one's self-evaluation. Therefore, a self-aware individual is more attentive to other's perceptions, which results in a more accurate self-rating. One study has shown that raters high in self-awareness were more accurate in their self-evaluations (Atwater & Yammarino, 1992). Gibbons (1983) demonstrated that self-ratings correlated more highly with performance measures for raters high in private self-consciousness and internal focus of attention.

The egocentric bias theory, self-enhancement theory, self-presentation theory, and self-awareness theory have considerable overlap and mutually affect self-other discrepancies. In the case of MSF, self-awareness theory plays a very significant role. MSF is primarily used for developmental purposes and is anonymous to promote honest feedback (Guest & Blucher, 1998). It would not seem advantageous for leaders to purposefully enhance their self-ratings, as the results do not dictate promotions or pay but are primarily developmental in purpose. Additionally, the process is anonymous and rarely does anyone other than the target leader see the results. It is also generally understood that self-ratings are separated from the aggregate total. Thus, it would seem that the discrepancy between self-ratings and other's ratings may be primarily due to poor self-awareness. Self-awareness theory has been adopted as the primary

explanation for self-other agreement in the current literature, and self-other agreement has been operationalized as a level of one's self-awareness (Atwater & Brett, 2006; Fleenor, et al., 1996).

Transformational Leadership Theory

Self-awareness and understanding follower's perceptions are strongly associated with two of the four transformational leadership dimensions in transformational leadership theory (Bass, 1985). Transformational leadership encompasses four dimensions, which include inspirational motivation, individualized consideration, intellectual stimulation, and idealized influence (charisma) (Avolio, Waldman, & Yammarino, 1991; Bass, 1990).

Inspirational motivation is the degree to which leaders inspire followers with a strong vision of the future. Inspirational motivation includes clear communication of expectations and motivation through goals based on a shared vision. This aspect of transformational leadership involves alignment of perceived vision with followers. In theory, those with high self-awareness are better able to share a common vision based on an understanding of follower's perception and the leader's self-perception.

Individualized consideration is the degree to which leaders understand the individual needs of their followers, and develop individuals through coaching. This is accomplished by establishing a supportive climate, as well as providing opportunities for growth. Self-awareness is needed to understand how one's behaviors are affecting others, as well as being able to recognize the developmental needs in others.

The other two dimensions of transformational leadership less associated with self-awareness are idealized influence and intellectual stimulation. *Idealized influence (charisma)* is the degree to which leaders influence others by example. These charismatic leaders appeal to people on an emotional level and display conviction. Idealized influence involves trust, admiration, and respect. Followers identify with the leader and wish to emulate them. *Intellectual*

stimulation is the degree to which leaders encourage innovation and divergent thinking. In this dimension leaders create a climate of creativity by challenging norms and taking calculated risks.

Also part of the transformational leadership theory is a style of leadership called transactional leadership. Different from transformational leadership, transactional leadership is based on a simple exchange relationship with followers (Bass & Avolio, 1994). For example, transactional leaders will specify job requirements and set conditions for an exchange, such as pay or compensation. Transactional leadership has two dimensions, contingent reward and management by exception-active. *Contingent reward* is the degree to which leaders establish productive transactions with followers. Leaders high in contingent reward clearly indicate expectations and establish appropriate rewards for meeting those expectations. *Management by exception-active* is the degree to which leaders take corrective action when followers fail to meet the expectations of a leader-follower exchange. Generally, leaders scoring high in management by exception-active tend to be less involved when followers are performing as expected. Transactional leadership is not antipodal to transformational leadership, but rather augments transformational leadership (Bass, 1990; Judge & Piccolo, 2004). This means that those high in transformational leadership tend to be effective transactional leaders as well.

A third style of leadership in the transformational leadership theory is non-leadership. Non-leadership has two dimensions, which include management by exception-passive and laissez-faire leadership. *Management by exception - passive* is similar to management by exception-active in that these leaders do not intervene until followers fail to perform as expected. Management by exception-active and management by exception-passive differ in the timing of the leader's intervention. Active leaders anticipate problems and monitor follower behavior to take corrective action before problems become serious. Passive leaders wait until problems are

serious before taking corrective action. *Laissez-faire leadership* is the absence or avoidance of leadership. These leaders avoid responsibility, are disorganized, and offer little in terms of direction or support. Although laissez-faire leadership may seem similar to management by exception-passive, it should be treated as a different dimension because it represents the absence of leadership (Bass, 1998). Transformational leadership theory originally categorized management by exception-passive with the transactional leadership style. However, due to factor analysis results management by exception-passive has been included with laissez-faire leadership to form the non-leadership style (Avolio & Bass, 2004).

Transformational leaders have been shown to be more effective than transactional leaders because of their ability to promote purpose and vision, which transcends short term exchanges. Empirical results have shown that transformational leaders have helped organizations increase production, group effectiveness, organizational sales, market shares, and employee moral (Bass, 1998; Bass & Avolio, 1994; Geyer & Steyrer, 1998). Followers of transformational leaders were found to have higher job satisfaction, motivation, and more positive emotions through the workday than transactional leaders (Bono, Foldes, Vinson, & Muros, 2007; Judge & Piccolo, 2004). Transformational leaders, when compared to transactional leaders, received higher performance appraisals, higher effectiveness ratings, and higher ratings on objective performance measures (Geyer & Steyrer, 1998; Hater & Bass, 1988). Judge and Piccolo (2004) conducted a meta-analysis and found high correlations of transformational leadership and follower satisfaction, follower motivation, and leader effectiveness (r = .53 - .71), and high negative correlations with laissez-faire leadership and follower satisfaction and leader effectiveness (r = -.54 and -.58 respectively). Also, manager's transformational leadership score was found to be positively related to employee performance and customer loyalty (Liao & Chuang, 2007).

Transformational leaders are more concerned with the view point of their staff than transactional leaders, and more likely to solicit feedback (Alimo-Metcalfe, 1998). Leaders who were in agreement with other's ratings when measuring leadership behavior (self-aware of their management style) were also found to be more charismatic in their leadership style (Sosik, 2001). A leader is more likely to be seen as transformational if there is a strong relationship between a leader's self-perception and the perception of their employees (Alimo-Metcalfe, 1998; Bass & Yammarino, 1991).

In theory, transformational leaders are self-aware leaders. Transformational leaders align individuals with organizational values and goals (Sosik, 2001). This alignment is best achieved when leaders are self-aware; meaning that they are able to recognize the needs of followers, see them self as others do, and adjust their behavior in order to get followers to align with organizational goals and values (Sosik, 2001). It is meaningful to measure a leader's level of self-awareness because self-awareness is a critical element of transformational leadership.

Research Objectives

The purpose of this study is to explore the usefulness of self-other discrepancies in MSF and its relationship with transformational leadership. One of the difficulties in interpreting MSF is making sense of the discrepancies that exist between a manger's self-ratings and other's ratings on competency skills. Does awareness of other's perceptions make one a better leader? The literature suggests that agreement between a manager's self-ratings and ratings from other sources can predict leadership style. It is hypothesized in this study that self-other agreement can predict transformational leadership.

This study expands the current literature in two ways. First, current research in self-other agreement has not explored implications of transformational leadership, specifically, that those high in self-other agreement will be more likely to have qualities of a transformational leader.

Secondly, this study will examine alternative techniques in measuring self-other agreement that are more pragmatic for organizations using an MSF process. Findings from this study can help companies in giving more appropriate feedback, and assist leaders in making sense of the self-other discrepancies that commonly occur in MSF. Also, by making the connection between self-other agreement and transformational leadership organizations can better their efforts of developing their own leaders.

Hypotheses

H1: It is hypothesized that a leader high in self-other agreement will be seen as a more transformational leader.

It is hypothesized that leader's self-other ratings high in agreement will be positively related to inspirational motivation and individualized consideration as measured by the Multi-factor Leadership Questionnaire (MLQ). That is, as self ratings and other's ratings have high concordance, then inspirational motivation and individualized consideration will also be the highest. Since the sub-scales of transformational leadership are highly correlated (Judge & Piccolo, 2004), it is hypothesized that the overall transformational leadership score will also be positively related to self-other agreement. Conversely, management by exception-passive and laissez-faire leadership, the measurement of non-leadership, will be highest for those lower in self-other agreement.

H2: It is hypothesized that a leader with direct reports high in agreement will be seen as a more transformational leader.

Transformational leaders are able to create a consensus, or agreement among followers (Sosik, 2001). Also, transformational leaders are more likely to have followers who agree in their perceptions of the leader (Feinberg, Ostroff, & Burke, 2005). It is hypothesized that a leader with direct reports high in agreement will be seen as a more transformational leader. Conversely,

management by exception-passive and laissez-faire leadership will be highest for managers with direct reports low in agreement.

H3: It is hypothesized that a leader high in self-other agreement and that has direct reports high in agreement will be seen as a more transformational leader.

The previous hypotheses will come together to form the final hypothesis that a leader high in self-other agreement and that has direct reports high in agreement will have higher transformational leadership ratings. Also, a leader that has low self-other agreement but high direct report agreement will have low transformational leadership ratings, and high management by exception-passive and laissez-faire leadership ratings.

METHOD

Participants

Participants in this study were mid-level and upper-level executives from four companies located within the United States. Data were collected as part of an MSF process directed by the Human Resources (HR) manager of each company. A total of 34 leaders and 254 raters participated in this study. Raters consisted of the leaders' managers, peers and direct reports. Three leaders and 21 raters were excluded from the analysis because of incomplete surveys, leaving 31 leaders and their 233 respective raters included in the analysis. The mean number of raters for each leader was 7.52 (*SD* = 3.29), with an average of 1.03 supervisor ratings, (*SD* = 0.85), 2.45 peer ratings (*SD* = 1.77), and 4.03 (*SD* = 2.80) direct report ratings. Of the 31 leaders, 25 were male and 6 were female.

Setting

A total of four companies participated in this study. Three companies from the Mountain West were invited to participate in the current study at a monthly luncheon for HR managers. Another company from the Northeast was referred by an HR manager familiar with the study. Two of the companies provided Information Technology (IT) services, while the other two companies were in the manufacturing or financial services industry. A summary of demographic variables is shown in Table 1.

Measurements

Multi-Source Feedback Survey. The beginning part of the MSF survey contained an informed consent form, an introduction to MSF, and basic demographic information (see Appendix A). The raters were asked for the name of the leader being rated as well as their relationship to the leader (e.g. peer, direct report). To assure that the raters had adequate opportunities to assess the leader at work, two questions asked how long the rater had *known* the

Table 1

Demographic Information of the Four Participating Companies

	Company 1	Company 2	Company 3	Company 4	Totals
Location	Mountain West	Mountain West	Northeast	Mountain West	
Industry	Information Technology	Manufacturing	Information Technology	Financial Services	
Leaders	11 (35.5%)	6 (19.4%)	13 (41.9%)	1 (3.2%)	31 (100%)
Raters	82 (35.2%)	40 (17.2%)	104 (44.6%)	7 (3.0%)	233 (100%)

leader, and how long the rater had *worked* with the leader. An additional question had the raters identify on a 5-point scale how well they we able to observe the target leader's behavior at work.

To assess self-other agreement and direct report agreement, personalized MSF surveys were designed for each participating company (see Appendix B). All participants within the same company completed the same survey. It was not necessary for the MSF survey to be uniform across companies because the research question focused on self-other agreement on each item, rather than competency totals. Although the surveys were personalized for each company, many of the same questions were used, and each survey was relatively similar in length. The MSF surveys designed for the companies contained 50 to 57 items or behavioral statements. Each behavioral statement was rated on a 5-point scale ranging from (1) *never* to (5) *always*. The MSF survey also include the option of *no opportunity to assess*, and was described as "not having the ability to assess this behavior at work." The *no opportunity to assess* option is essential in the multi-source feedback process as the participants will simply guess if they are not provided with this item (Atkins & Wood, 2002). Responses for *no opportunity to assess* were excluded from the analysis. Items for the self-assessment differed slightly from the other rater's

survey. For example, the question "The person I am rating… identifies the core element of an issue," will read "I identify the core element of an issue" on the self-assessment. MSF surveys also contained qualitative questions that were not used in this study, but were valuable for feedback purposes.

Self-Awareness Questionnaire. A set of four questionnaire items were used as a direct measure of self-awareness. Despite the difficulties in measuring self-awareness, instruments have been developed for others to rate a leader's level of self-awareness. The current study used the measure by Van Velsor et al. (1993) due to its relative ease to administer and simplicity (Fleenor, et al., 1996). The questionnaire has four items to be completed by those familiar to the leader. The items were included in each company's MSF survey and used the same 5-point scale. The four items were "The person I am rating … (1) Is aware of the impact of his/her behavior on others, (2) Has an accurate picture of his/her strengths and weaknesses, (3) Attempts to do things unsuitable to his/her skill level (this item is reverse scored), and (4) Is sensitive and aware of what others think of him/her.

Multi-Factor Leadership Questionnaire (MLQ). The MLQ Form 5X (Bass & Avolio, 1994) was used to assess the transformational leadership, transactional leadership and non-leadership behaviors of the participating leaders. The MLQ Form 5X contains 45 questions using a 5-point scale ranging from (0) *not at all* to (4) *frequently, if not always*. However, the scale was converted to a 1 - 5 scale rather than a 0 - 4 scale in order to be consistent with other items on the MSF survey. Items are worded the same for direct reports, peers and supervisors.

Transformational leadership has four dimensions: inspirational motivation, individualized consideration, intellectual stimulation and idealized influence (charisma). Four items on the MLQ measure each dimension. Sample items from each subscale include (a) inspirational

motivation, “talks optimistically about the future,” (b) individualized consideration, “gives personal attention to others when necessary,” (c) intellectual stimulation, “shows others how to think about problems in new ways,” and (d) idealized influence, “displays a sense of power and influence.” Transaction leadership also has four items for each of the two dimensions: contingent reward and management by exception-active. Non-leadership, which is an absence of leadership behavior, was measured by the dimensions of management by exception-passive and laissez-faire leadership, each consisting of four items. The remaining nine questions measured leadership satisfaction and leadership effectiveness, which were not relevant to the current study and not included in the survey.

Procedure

The researcher coordinated the data collection through each company’s HR managers. Meetings were held with each HR manager, along with other leaders of the organization, to develop an MSF survey. Competencies and individual items were chosen based on the organizations strategic goals and mission statement. The MSF survey, along with the MLQ and self-awareness items, was complied into an on-line survey administered through surveymonkey.com. Each company gave consent to participate in the study as well as each individual rater. Raters had the option of withdrawing without penalty. The rater’s responses were also kept anonymous to ensure candid and helpful feedback.

The HR manager of each company identified the target leader’s supervisor, peers and direct reports that had worked with the leader long enough to observe that leader’s behavior in a variety of settings. Instructions and survey links were sent to each rater through their company e-mail. Each rater completed the survey on a personal computer at work, rating their assigned leader’s competency skills, self-awareness and leadership style. Additionally, each target leader

completed a self-assessment of the same survey. The survey took approximately 20 to 30 minutes to complete.

Feedback reports were provided for each manager that participated in the study. Each report contained five sections. The first section was an introduction to MSF, and an explanation of the types of raters involved in the feedback process. The second section contained an item by item analysis, displaying the self-assessment score, other's mean score, and the company norm (the aggregate of all the raters in the company for each item). Data were aggregated across raters to form a single score for each item to ensure confidentiality. Other's mean scores were broken down by peer and direct report means in cases where there were more than three raters in each category. The third section contained the four self-awareness items and was displayed in the same format as section two. The fourth section summarized the target leader's highest scores (strengths), lowest scores (weaknesses), and self-other discrepancies (gaps). The final section contained the transformational leadership, transactional leadership and non-leadership scores. Due to copyright laws the individual items of the MLQ were not shown in the reports. However, a total aggregate was displayed for each leadership dimension, along with a brief explanation of transformational leadership theory. HR managers used the feedback reports for goal setting sessions, assessing training needs and leadership development purposes. The feedback was purely for developmental purposes and was not used in any promotion, pay or bonus decisions.

Data Analysis

The direct report, peer and supervisor ratings on the MSF survey were aggregated across each item. This mean score on each item was compared to the target leader's self-ratings. These two ratings were used to calculate a reliability coefficient, which will be called *self-other agreement r*. The reliability coefficient was calculated by dividing the mean covariance by the mean variance from the self-other variance-covariance matrix. A mean difference squared score

was also calculated for each target leader as an alternative self-other agreement measurement, which will be called *self-other agreement* D^2. This was calculated by subtracting the mean other score from the self-rating on each item. The difference score was then squared in order to display the total self-other deviation (both over rating and under rating). Since the MSF surveys differed slightly between each company the difference squared score was aggregated across each item.

A mean self-awareness score was calculated for each manager by aggregating each item on the self-awareness scale. The item, "attempts to do things unsuitable to his/her skill level" was reverse scored before aggregates were calculated. As the dependant measure a transformational leadership, transactional leadership and non-leadership score, as well as a score for each leadership dimension, was calculated for each leader.

It was hypothesized that a leader high in self-other agreement will be seen as a more transformational leader. A regression analysis will be used to test the first hypothesis, with the *self-other agreement r* as the predictor variable and the dependant variable being the transformational leadership scores. The *self-other agreement* D^2 score will also be used an alternative measurement for self-other agreement. Because of the difficulties of operationalizing self-other agreement as self-awareness, the direct measure of self-awareness will also be used to test the first hypothesis. Regression analysis will be used to test if this measure of self-awareness predicts transformational leadership.

It was also hypothesized that a leader with direct reports high in agreement will be seen as a more transformational leader. The self-assessment and the direct report assessment scores will be used to calculate each leader's direct report agreement. The measure of agreement will be calculated using the same reliability coefficient of the average covariance divided by the average variance. The reasoning for using only the direct report's rating for this measure of agreement is

that transformational leadership implies a leader-follower relationship, not a leader-peer or leader-supervisor relationship. A regression analysis will be used to test the second hypothesis, with the direct report's agreement coefficient as the predictor variable and the leader's transformational leadership score as the dependant variable.

The third hypothesis stated that a leader high in both self-other agreement and direct reports agreement will be seen as a more transformational leader. To test the third hypothesis a multiple regression analysis will be used, with the *self-other agreement r* and the direct report's reliability coefficient as the predictor variables and the leader's transformational leadership score as the dependant variable. A three dimensional graph will also be created to determine the shape of the relationship between manager's self-other agreement (X), direct report agreement (Y), and transformational leadership scores (Z).

RESULTS

Due to the vast differences in company culture and leadership style (Burns, 1979), an analysis of variance (ANOVA) was conducted to find any individual differences among companies. The four companies in this study did not differ significantly from one another on any of the variables (see Table 2). The *F* values were all non-significant and ranged from 0.10 (*self-other agreement r*) to 2.09 (management by exception-passive). These results indicate that each company had relatively similar scores on each of the predictor variables and transformational leadership scores.

Raters were asked how long they have known their leader and how long they have worked with their leader (see Appendix A). It is suggested that raters have enough time to observe their leader's behavior but not long enough that the raters develop a leniency bias (Morgeson, Mumford, & Campion, 2005; Moser, et al., 1999). The mean number of months raters have known their leader was 35.65 (*SD* = 40.23), and the mean number of months they have worked with their leader was 31.22 (*SD* = 34.85). Meaning, the average rater has known their leader for about three years and has worked with them just over two and a half years, which is within the appropriate range as suggested by Moser, et al.

Raters were also asked how well they have been able to observe their leader's behavior at work. All the raters reported some regular contact with their leader, with the rater's average contact being a 4.06 (*SD* = 0.84) on a 5-point scale. When reporting the amount of contact with their leader, 91 (36.1%) raters indicated a 5 (I interact daily with this person in a variety of settings) and 90 (35.7%) raters indicated a 4 (I interact with this person almost every day in more than one setting). These results suggest that the raters had enough contact with their leaders to accurately report their behavior (see Table 3).

Table 2

Analysis of Variance for Company Variables

Variables	Company 1	Company 2	Company 3	Company 4		
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*
Self-Other Agreement r	.29 (.18)	.33 (.19)	.29 (.15)	.28*	0.10	.96
Self-Other Agreement D^2	.66 (.33)	.61 (.40)	.53 (.26)	.49*	0.40	.75
Self-Awareness	3.91 (.28)	4.06 (.37)	4.02 (.58)	3.87*	0.20	.90
Direct Report Agreement	.20 (.15)	.26 (.13)	.11 (.16)	.02*	1.81	.17
Transformational Leadership	3.88 (.38)	4.03 (.27)	4.04 (.36)	4.19 (.78)	0.61	.61
Idealized Influence	3.89 (.40)	4.06 (.31)	4.06 (.46)	4.26 (.68)	0.61	.61
Inspirational Motivation	3.96 (.45)	3.98 (.30)	4.13 (.41)	4.24 (.86)	0.48	.70
Intellectual Stimulation	3.74 (.40)	3.99 (.31)	3.98 (.39)	4.17 (.75)	1.19	.33
Individualized Consideration	3.91 (.46)	4.08 (.35)	3.98 (.37)	4.09 (.84)	0.26	.85
Transactional Leadership	3.33 (.40)	3.51 (.27)	3.50 (.16)	3.82 (1.03)	1.31	.29
Contingent Reward	3.88 (.38)	3.90 (.20)	4.10 (.36)	4.28 (.60)	1.34	.28
MBE-active	2.79 (.71)	3.11 (.48)	2.90 (.32)	3.37 (1.46)	0.80	.50
Non-Leadership	1.79 (.31)	1.66 (.32)	1.92 (.47)	1.45 (.03)	1.35	.28
MBE-Passive	2.07 (.37)	1.73 (.26)	2.22 (.56)	1.69 (.23)	2.09	.13
Laissez-Faire Leadership	1.51 (.27)	1.58 (.43)	1.62 (.39)	1.20 (.16)	0.88	.46
Number of Raters	7.82 (3.16)	6.83 (2.23)	8.23 (3.86)	6.00 (1.41)	0.43	.73

Note. N = 32

*N=31, *n* = 1 data was only available for one leader

Table 3

Mean and Standard Deviation Scores for the Background Questionnaire

Questionnaire items	N	*M*	*SD*
Number of months raters have known their leader	231	35.65	40.23[a]
Number of months raters have worked with their leader	232	31.22	34.95[b]
Rater's ability to observe leader's behavior (5-point scale)	233	4.06	0.84
Total number of raters (raters per leader)	233	7.52	3.29
Supervisors (supervisors per leader)	32	1.03	0.85
Peers (peers per leader)	76	2.45	1.77
Direct Reports (direct reports per leader)	125	4.03	2.80

[a] Range = 2 to 272 months. [b] Range = 1 to 272 months.

Table 4

Descriptive Statistics and t Values for Transformational Leadership, Transactional Leadership and Non-Leadership

	M	*SD*	Norms	*t*	*p*
Transformational Leadership	4.00	.37			
Idealized Influence	4.02	.42	4.00	.31	.76
Inspirational Motivation	4.07	.42	4.00	.87	.39
Intellectual Stimulation	3.92	.40	3.75	2.34*	.03
Individualized Consideration	4.00	.41	4.00	-.06	.95
Transactional Leadership	3.48	.35			
Contingent Reward	4.01	.37	4.00	.09	.93
MBE-Active	2.95	.58	2.67	2.69*	.01
Non-Leadership	1.81	.38			
MBE-Passive	2.05	.47	2.00	.60	.55
Laissez-Faire Leadership	1.56	.35	1.50	.98	.34

Note. N = 31

* $p < .05$

Descriptive statistics for transformational leadership are shown in Table 4. Along with transformational leadership means and standard deviations, national averages (norms) for each dimension are also shown (provided by the MLQ manual; Avolio & Bass, 2004). The sample for the national average consisted of 27, 285 ratings of 3,755 leaders within the United States. A one-sample t-test was conducted to see if the leaders' transformational leadership score in the current study differed from the national average. Leaders in the current study did not differ significantly from the national average, with the exception of intellectual stimulation and management by exception-active. Leaders in this study scored higher on intellectual stimulation and management by exception-active.

Results thus far indicate that raters had sufficient observational experience to rate their leaders accurately, that leaders did not differ significantly from other leaders in the United States, and that the four participating companies did not significantly differ from each other. These results also indicate that the MSF process was done appropriately, which strengthens internal validity, and that the transformational leadership scores were within the range of other companies, which strengthens external validity.

Hypothesis 1

It was hypothesized that self-other agreement would significantly predict transformational leadership, specifically the dimensions of inspirational motivation and individualized consideration. To test this hypothesis *self-other agreement r* was used as the predictor variable in a regression analysis, with the dimensions of transformational leadership as the dependent variable. Results indicate that *self-other agreement r* does not significantly predict any dimension of transformational leadership. Inspirational motivation ($R^2 = .00$, F (1, 29) = 0.04, $p = .85$), individualized consideration ($R^2 = .01$, F (1, 29) = 0.25, $p = .62$), and overall

transformational leadership (R^2 = .00, F (1, 29) = 0.01, p = .91), did not have a significant proportion of variance accounted for by self-other agreement (see Table 5).

It was also hypothesized that management by exception-passive and laissez-faire leadership could be predicted based on a negative relationship with self-other agreement. Management by exception-passive and laissez-faire leadership was not significantly predicted from self-other agreement. Thus, *self-other agreement r* did not significantly predict dimensions of transformational leadership or dimensions of non-leadership.

As an alternative measurement, *self-other agreement* D^2 or the aggregate difference squared scores, were used to predict transformational leadership. *Self-other agreement* D^2 significantly predicted transformational leadership and individualized consideration. *Self-other agreement* D^2 accounts for 16% of the variance in transformational leadership and 17 % of the variance in individualized consideration. The relationship between *self-other agreement* D^2 and transformational leadership is negative because lower difference scores indicate a higher level of self-other agreement. *Self-other agreement* D^2, however, did not significantly predict inspirational motivation or the dimensions of non-leadership. Results for the other dimensions of transformational leadership are shown in Table 6.

The direct measure of self-awareness consisted of four items aggregated together to form an index of self-awareness for each leader. This index of self-awareness significantly predicted several dimensions of transformational leadership (see Table 7). In line with hypothesis one, self-awareness significantly predicted transformational leadership, individualized consideration and inspirational motivation. Additionally, self-awareness significantly predicted the dimensions of non-leadership through a negative relationship.

Hypothesis 2

Transformational leaders are able to create a consensus, or agreement among followers, and are more likely to have followers who agree in their perceptions of the leader (Feinberg, et al., 2005; Sosik, 2001). It was hypothesized that a leader with direct reports high in agreement will be seen as a more transformational leader. Conversely, management by exception-passive and laissez-faire leadership will be highest for managers with direct reports low in agreement. The second hypothesis was not supported by the regression analysis (see Table 8). Direct report agreement did not significantly predict any dimension of transformational leadership, transactional leadership or non-leadership. Direct report agreement only explained a marginal amount of transformational leadership, and best predicted inspirational motivation (explaining only 3% of the variance).

Table 5

Regression Analysis for Self-Other Agreement r and Transformational Leadership, Transactional Leadership and Non-Leadership

Leadership Dimensions	R^2	*B*	SE *B*	ß	*t*	*p*
Transformational Leadership	.00	.05	.42	.02	0.12	.91
Idealized Influence	.00	.00	.47	.00	-0.01	.99
Inspirational Motivation	.00	-.09	.48	-.04	-0.19	.85
Intellectual Stimulation	.00	.07	.46	.03	0.14	.89
Individualized Consideration	.01	.23	.46	.09	0.50	.62
Transactional Leadership	.00	-.10	.40	-.05	-0.24	.81
Contingent Reward	.01	-.26	.41	-.12	-0.62	.54
MBE-Active	.00	.07	.65	.02	0.10	.92
Non-Leadership	.02	.28	.42	.12	0.67	.51
MBE-Passive	.01	.30	.53	.10	0.56	.58
Laissez-Faire Leadership	.02	.27	.39	.13	0.69	.50

Note. N = 31

Table 6

Regression Analysis for Self-Other Agreement D^2 and Transformational Leadership, Transactional Leadership and Non-Leadership

Leadership Dimensions	R^2	*B*	SE *B*	ß	*t*	*p*
Transformational Leadership	.16	-.49	.21	-.40	-2.38*	.02
Idealized Influence	.19	-.60	.23	-.44	-2.62*	.01
Inspirational Motivation	.09	-.42	.24	-.31	-1.73	.09
Intellectual Stimulation	.09	-.40	.23	-.31	-1.72	.10
Individualized Consideration	.17	-.55	.22	-.41	-2.44*	.02
Transactional Leadership	.00	-.04	.21	-.03	-0.17	.87
Contingent Reward	.12	-.42	.21	-.35	-2.00	.05
MBE-Active	.03	.34	.34	.18	1.00	.33
Non-Leadership	.08	.34	.22	.28	1.55	.13
MBE-Passive	.08	.45	.27	.29	1.63	.11
Laissez-Faire Leadership	.04	.24	.21	.21	1.15	.26

Note. N = 31

* $p < .05$

Table 7

Regression Analysis for Self-Awareness and Transformational Leadership, Transactional Leadership and Non-Leadership

Leadership Dimensions	R^2	*B*	SE *B*	ß	*t*	*p*
Transformational Leadership	.31	.48	.13	.56	3.61***	.00
Idealized Influence	.40	.62	.14	.64	4.42***	.00
Inspirational Motivation	.19	.43	.16	.44	2.63*	.01
Intellectual Stimulation	.29	.51	.15	.54	3.47**	.00
Individualized Consideration	.16	.37	.16	.39	2.31*	.03
Transactional Leadership	.03	.13	.15	.16	0.87	.39
Contingent Reward	.24	.41	.14	.49	3.01**	.01
MBE-Active	.01	-.16	.25	-.12	-0.63	.54
Non-Leadership	.31	-.49	.14	-.56	-3.60***	.00
MBE-Passive	.28	-.57	.17	-.52	-3.31**	.00
Laissez-Faire Leadership	.25	-.40	.13	-.50	-3.09**	.00

Note. N = 31

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

Table 8

Regression Analysis for Direct Report Agreement and Transformational Leadership, Transactional Leadership and Non-Leadership

Leadership Dimensions	R^2	*B*	SE *B*	ß	*t*	*p*
Transformational Leadership	.00	-.14	.44	-.06	-0.31	.76
Idealized Influence	.01	-.26	.49	-.10	-0.52	.61
Inspirational Motivation	.03	-.46	.49	-.17	-0.94	.36
Intellectual Stimulation	.00	-.06	.48	-.02	-0.12	.91
Individualized Consideration	.01	.23	.48	.09	0.48	.63
Transactional Leadership	.00	-.08	.41	-.04	-0.20	.85
Contingent Reward	.03	-.42	.42	-.18	-0.99	.33
MBE-Active	.01	.26	.68	.07	0.38	.71
Non-Leadership	.01	.28	.44	.12	0.64	.53
MBE-Passive	.00	.20	.56	.07	0.25	.73
Laissez-Faire Leadership	.03	.37	.40	.17	0.92	.36

Note. N = 31

Table 9

Multiple Regression Analysis for Self-Other Agreement and Direct Report Agreement Predicting Transformational Leadership, Transactional Leadership and Non-Leadership

		Self-other Agreement			Direct Report Agreement		
Leadership Dimensions	R^2	*B*	SE *B*	ß	*B*	SE *B*	ß
Transformational Leadership	.01	.14	.48	.06	-.20	.50	-.08
Idealized Influence	.01	.13	.54	.05	-.32	.56	-.12
Inspirational Motivation	.03	.14	.53	.05	-.52	.56	-.20
Intellectual Stimulation	.00	.11	.52	.05	-.11	.54	-.04
Individualized Consideration	.01	.17	.52	.07	.15	.54	.06
Transactional Leadership	.00	-.08	.45	-.04	-.05	.47	-.02
Contingent Reward	.03	-.09	.47	-.04	-.38	.48	-.16
MBE-Active	.01	-.06	.74	-.02	.29	.77	.08
Non-Leadership	.02	.20	.48	.09	.19	.50	.08
MBE-Passive	.01	.27	.61	.09	.07	.63	.02
Laissez-Faire Leadership	.03	.14	.44	.06	.31	.46	.14

Note. N = 31

Hypothesis 3

The previous two hypotheses formed the final hypothesis that a leader high in self-other agreement and that have direct reports high in agreement will be seen as a more transformational leader. Also, those that have low self-other agreement but high direct report agreement will have low transformational leadership scores, and high management by exception-passive and laissez-faire leadership scores. Multiple regression results do not support hypothesis three.

Self-other agreement r and direct report agreement only accounted for a marginal percent of the variance in transformational leadership scores (see Table 9). The largest R^2 was for inspirational motivation (R^2 = .03, F (1, 29) = 0.46, p = .64) and laissez-faire leadership (R^2 = .03, F (1, 29) = 0.46, p = .64). Neither self-other agreement nor direct report agreement significantly predicted transformational leadership. Standardized betas were non-significant, ranging from -.16 (p = .51) to .14 (p = .44).

To visually explore the relationship between leaders' self-other agreement, direct report agreement and transformational leadership, a 3-dimensional graph was created. The three dimensional graph can be seen in Figure 3. Since the multiple regression analysis was non-significant, no other significant findings from the graph can be detected. No patterns involving the three variables can be deciphered by visual inspection of the graph. A possible explanation for the non-significant findings is that leaders higher in transformational leadership are equally distributed across self-other agreement. Leaders high in transformational leadership have a tendency to be lower on direct report agreement. However, most of the leaders have a tendency to be lower on direct report agreement, regardless of their transformational leadership score.

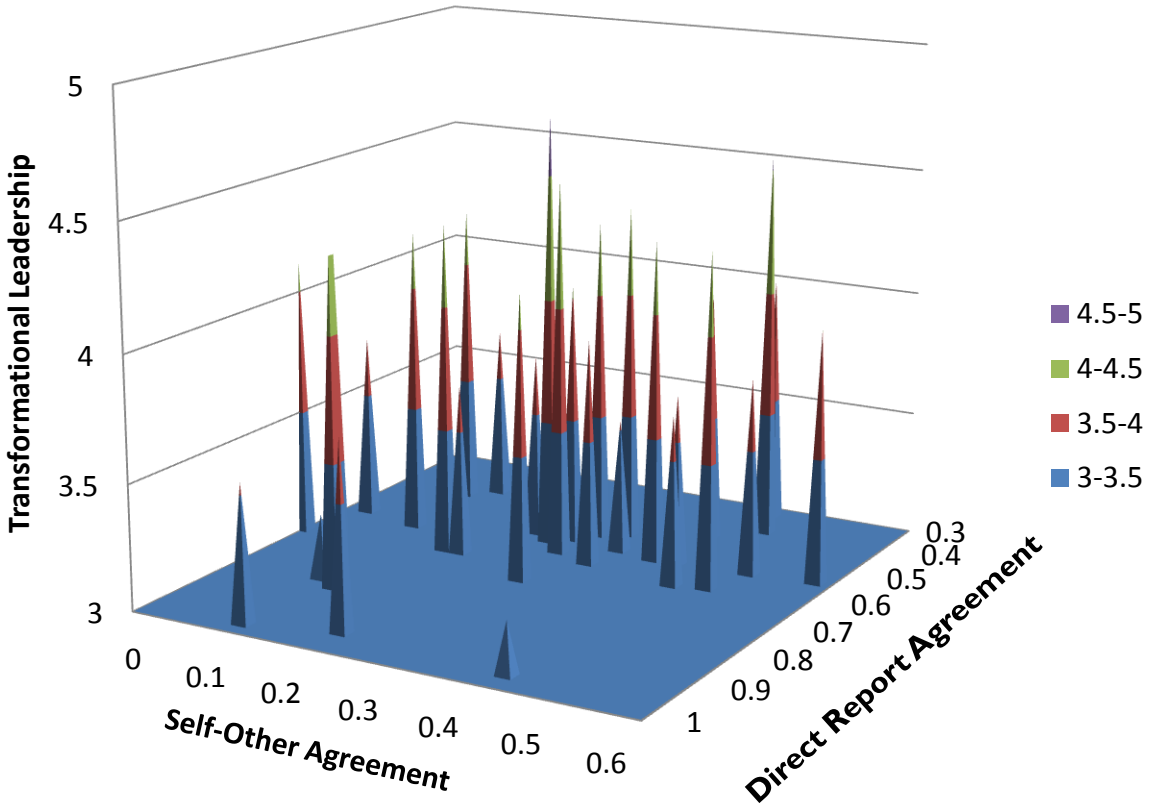

Figure 3. A 3-dimensional graph representing the relationship between leaders' self-other agreement, direct report agreement and transformational leadership.

Additional Findings

Gender Differences. Gender differences among leaders were examined in addition to the initial hypotheses. There were no significant differences between male and female leaders for any measure self-other agreement. Both male (M = .30, SD = .15) and female (M = .29, SD = .23) leaders did not differ significantly when measuring *self-other agreement r* [t (29) = 0.07, p = .95]. Male (M = .62, SD = .33) and female (M = .47, SD = .16) leaders also did not differ in *self-other agreement* D^2 [t (29) = 1.04, p = .31]. When measuring direct report agreement there were also no differences between male (M = .17, SD = .15) and female (M = .13, SD = .20) leaders [t (29) = 0.54, p = .59]. Both male (M = 3.95, SD = .42) and female (M = 4.12, SD = .51) leaders also did not differ significantly in their ratings of self-awareness [t (29) = -0.86, p = .40].

However, there were significant gender differences in leaders' transformational leadership score (see Table 10). Female leaders were perceived as more transformational than their male counterparts. When considering the individual dimensions of transformational and transactional leadership, females were rated higher than males on idealized influence, individualized consideration and contingent reward. In the non-leadership category male leaders were perceived to exhibit more management by exception-passive behaviors than female leaders. There were large effects sizes (d = 1.00 - 1.31) for each of the dimensions of transformational leadership that were significant (Cohen & Kazdin, 2003).

Self-Other Agreement and Self-Awareness. Correlations were calculated to find the relationship between the various independent variables. None of the independent variables had significant correlations. *Self-other agreement r* (r = -.20, p = .27) and *self-other agreement* D^2 (r = -.30, p = .10) did not significantly correlate with the direct measure of self-awareness. Also, the two measurements of self-other agreement did not have a significant correlation with one another (r = -.33, p = .07).

Total MSF Scores. A regression analysis was calculated in order to find if the raters had a tendency to rate their leaders similarly across the MSF survey and the dimensions of leadership style. The overall MSF score correlated significantly with transformational leadership ($r = .77$, $p < .001$). In fact, the overall MSF score significantly predicted transformational leadership on nearly every dimension (see Table 11). Possible explanations for this finding will be discussed further.

Table 10

Results for Gender Differences in Transformational Leadership, Transactional Leadership and Non-Leadership

	Male		Female				
Leadership Dimensions	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*
Transformational Leadership	3.94	.36	4.27	.30	-2.06*	.05	1.00
Idealized Influence	3.94	.40	4.37	.31	-2.43*	.02	1.20
Inspirational Motivation	4.03	.42	4.24	.40	-1.10	.28	0.51
Intellectual Stimulation	3.87	.39	4.14	.42	-1.51	.14	0.67
Individualized Consideration	3.91	.39	4.33	.33	-2.44*	.02	1.16
Transactional Leadership	3.42	.29	3.72	.51	-1.97	.06	.72
Contingent Reward	3.93	.35	4.30	.30	-2.36*	.03	1.14
MBE-Active	2.90	.50	3.14	.86	-0.89	.38	.34
Non-Leadership	1.85	.38	1.60	.30	1.50	.14	.73
MBE-Passive	2.14	.48	1.67	.16	2.38*	.02	1.31
Laissez-Faire Leadership	1.57	.32	1.54	.46	0.20	.85	.08

Note. N = 31

* $p < .05$.

Table 11

Regression Analysis for Total MSF Score and Transformational Leadership, Transactional Leadership and Non-Leadership

Leadership Dimensions	R^2	*B*	SE *B*	ß	*t*	*p*
Transformational Leadership	.59	.62	.10	.77	6.39***	.00
Idealized Influence	.58	.55	.09	.76	6.33***	.00
Inspirational Motivation	.35	.43	.11	.60	3.99***	.00
Intellectual Stimulation	.62	.59	.09	.79	6.83***	.00
Individualized Consideration	.39	.46	.11	.62	4.29***	.00
Transactional Leadership	.04	.16	.16	.19	1.03	.31
Contingent Reward	.45	.55	.11	.67	4.85***	.00
MBE-Active	.04	-.10	.10	-.20	-1.08	.29
Non-Leadership	.41	-.51	.11	-.64	-4.48***	.00
MBE-Passive	.28	-.34	.10	-.53	-3.37**	.00
Laissez-Faire Leadership	.45	-.58	.12	-.67	-4.89***	.00

Note. N = 31

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

DISCUSSION

Overview of Hypotheses

The use of MSF has become an important strategy for organizations in developing their own leaders (London & Smither, 1995). It is necessary for leaders to have some awareness of how others perceive them in order to properly implement feedback (Delmhorst, 2006). Theoretically, leaders that are aware of what others think of them should be able to better interpret and implement feedback from co-workers. Leaders that are more self-aware should then become better leaders because of their ability to implement feedback to improve their leadership skills. MSF could facilitate the process of improving leaders' awareness of strengths and weaknesses, understanding the situations in which they can succeed, and better understand the perceptions of their direct reports. In fact, the current study seeks to find the connection between how well leaders recognize others' perceptions and their own leadership ability.

Transformational leadership has been chosen for the current study because it has shown to be an effective leadership style and fits nicely with the construct of self-awareness. Transformational leadership has been shown to be an effective style of leadership in increasing productivity, group effectiveness, organizational sales, market shares, and employee moral (Bass, 1998; Bass & Avolio, 1994; Geyer & Steyrer, 1998). Transformational leaders also tend to receive higher performance appraisals and higher ratings on objective performance measures than transactional leaders (Geyer & Steyrer, 1998; Hater & Bass, 1988). Also, the concept of self-awareness fits nicely with the theory of transformational leadership. Transformational leaders are self-aware leaders. Transformational leaders are able to create a consensus among followers through recognizing the needs of followers, seeing them self as others do, and adjusting their behavior in order to get followers to align with organizational goals and values (Sosik, 2001). Research has shown that leaders high in transformational leaders are more

concerned about the view point of their staff, and are more likely to solicit feedback (Alimo-Metcalfe, 1998).

The current study explored the relationship between self-other agreement in MSF and transformational leadership. It was hypothesized that self-other agreement and direct report agreement could predict transformational leadership. Each of the hypotheses will be discussed in turn.

H1: It is hypothesized that a leader high in self-other agreement will be seen as a more transformational leader.

The current study only gives partial support to the first hypothesis that self-other agreement can predict transformational leadership. Results from this study returns to the reoccurring question in the literature: How does one measure self-other agreement? A new approach to measuring self-other agreement was explored in the current study. The average covariance divided by the average variance was used to create a reliability coefficient of agreement between the leaders' self-ratings and others ratings, called *self-other agreement r*. This measure did not significantly predict any dimension of transformational leadership.

Difference scores, a more simplistic measure of self-other agreement, was also used to predict transformational leadership, called *self-other agreement* D^2. This measure of self-other agreement significantly predicted transformational leadership. Fletcher and Bailey (2003) conducted a comparison study between difference scores' (congruence-*d*) and correlation coefficients' (congruence-*r*) ability to predict leadership effectiveness in 104 target managers. Leadership effectiveness was not measured by an index of transformational leadership, but rather the managers' annual performance appraisals. Similar to the current findings, no significant relationship existed between congruence-*r* and managers' performance appraisals. However,

there was a significant relationship between congruence-*d* and managers' performance appraisals ($r = -.40$, $p < .001$). Fletcher and Bailey's research, along with the findings from the present study lend support that difference scores hold more power in predicting leadership effectiveness. Although a better predictor of leadership effectiveness, difference scores may have confounding methodological issues. The implications of using differences score will be discussed later.

H2: It is hypothesized that a leader with direct reports high in agreement will be seen as a more transformational leader.

The second hypothesis examined the relationship between direct report agreement and transformational leadership. Imbedded in the theory of transformational leadership is the idea that transformational leaders are able to create a consensus and shared vision among their followers (Bass, 1985). A small body of literature has supported this theory by finding that transformational leaders are more likely to have followers who agree in their perceptions (Feinberg, et al., 2005; Sosik, 2001). This study further examined this theory by hypothesizing that leaders will be perceived as a more transformational leader if their direct reports are high in agreement. The results indicate that direct report agreement did not significantly predict transformational leadership.

It is possible that direct reports may have an agreement with the leader's vision, but do not agree on the specific skill set of the leader, as measured by the MSF survey. It is important in the theory of transformational leadership to distinguish what the followers have a consensus about. Sosik (2001) has extended transformational leadership theory to mean that followers would have a consensus on ratings scales of leadership behavior. Although the findings by Sosik were significant, the current study did not support the notion that followers of transformational leaders would have a consensus on competency ratings of their leader.

H3: It is hypothesized that a leader high in self-other agreement and that has direct reports high in agreement will be seen as a more transformational leader.

Leaders that agree with others about their competency skills and that have direct reports that agree with one another would seem to be a more transformational leader because of their ability to understand their direct reports and create a consensus among them. Likewise leaders that do not agree with their direct reports, but their direct reports agree with one another would seem to be rated as a less transformational leader. This would seem to be the case because there is a consensus among followers but the leader is unaware of that consensus. These exploratory research questions have not been examined in the current literature. Results fail to support the third hypothesis, suggesting that self-other agreement and direct report agreement have little power in predicting transformational leadership.

Self-Other Agreement

For the past 15 years there have been debates in the literature about how to measure self-other agreement in MSF (Atwater & Brett, 2006; Atwater & Yammarino, 1992). The current trend is the use of polynomial regression as opposed to the confounding categorization approach (Atkins & Wood, 2002). Polynomial regression accounts for the self scores, other's scores and the outcome measure to be treated as three distinct constructs (Edwards, 1995). This allows one to see the form of agreement between self-ratings and other's ratings compared to another measure. Results of polynomial regression produce a three-dimensional graph which shows the numeric value of the dependant variable along the line of perfect self-other agreement.

When calculating polynomial regression a mean score on each competency is used, rather than the variability of each individual item. Thus, polynomial regression measures stereotype accuracy not differential accuracy, which is a more meaningful self-other agreement index (Cline & Richards, 1960). Stereotype accuracy is the ability to rate performance generally by using self-

other mean scores, while differential accuracy is the ability to accurately determine an individual's strengths and weaknesses comparatively by looking at the variability in the individual items (Bernieri, et al., 1994; Cline & Richards, 1960). A reliability coefficient, which measures the variability in individual items, better captures the pattern of agreement between a self-rater and other raters.

Another difficulty with polynomial regression is in the interpretation. It is a complex procedure that requires subjective visual inspections of a 3-dimensional graph. Additionally, determining the significance of the slopes of the graphs can be problematic (Edwards, 1995). Edwards and Perry (1993, p. 1577), the initial proponents for the use of polynomial regression, even confessed that "coefficients from polynomial regression equations are often difficult to interpret." With the complexities of interpretation in academic circles, how difficult and unpractical would it be to use polynomial regression in an applied setting? Research in applied settings should be more pragmatic, with methods that are easier to administer and more practical to apply. A more pragmatic methodology for measuring self-other agreement in MSF would be difference scores.

Difference Scores

There are shortcomings when using differences scores as an index of self-other agreement (Cronbach & Furby, 1970; Edwards, 1995). Difference scores treat positive and negative values the same, thus reducing the self-other agreement index to a single direction (Delmhorst, 2006). Thus, difference scores constrain the functional form of the relationship between self-other scores. It has been argued that other methods should treat self-ratings and other ratings as two separate constructs, rather than turning the self-other difference into a univariate model as with difference scores (Atkins & Wood, 2002). Another problem is that when difference scores are used to measure self-other agreement, the independent and dependent

variables share components (Delmhorst, 2006). For example, if other's ratings are used to compute self-other agreement on MSF and transformational leadership there would be a tendency for raters to evaluate both measures similarly. Since leaders have a tendency to over-rate their performance, they will be more accurate if other's rate them high on the MSF survey. If others rate a leader high on the MSF survey they also will likely rate the leader high on transformational leadership. Thus, difference scores can be confounding.

Although both reliability coefficients and difference scores are both referred as an index of self-other agreement, they seem to be a measurement of two completely different constructs. Results from Fletcher and Bailey's (2003) comparison study showed no congruence between the two measures (difference scores and reliability coefficients), indicating that they were measuring completely different constructs. In the current study, the two measurements of difference scores and reliability coefficients did not have a significant correlation, which supports previous research. Thus, there are many forms of self-other agreement, yet each measurement seems to capture a different construct.

Difference scores seem to be capturing the construct of *overall* differences between self and other ratings, regardless of direction, while the reliability coefficients seem to be capturing the *pattern* of agreement between self and other ratings or differential accuracy (Cline & Richards, 1960; Fletcher & Baldry, 2000). The difference scores better represent a common procedure used by consultants when providing MSF reports, referred to as the *gap analysis*. *Gaps* are the largest discrepancies between the self score and the average score from the other raters for each item. These gaps indicate where leaders and others disagree on ratings. There are no methodological problems with using difference scores item by item. However, difference scores start to have problems when they are used to calculate an overall index of agreement.

Although other measurements, such as polynomial regression or reliability coefficients, seem to be the more accurate measure of self-other agreement, they lend no practical value to the applied world. Few companies or HR manager will understand the process of calculating reliability coefficients, and the 3-dimensional graph produced from polynomial regression equations will be of little value to leaders. The current study supports the notion that difference scores, despite methodological problems, are more practical in applied settings. Additionally, difference scores hold more power in predicting effective leaders.

Self-Awareness

Self-awareness is a difficult construct to operationalize (Delmhorst, 2006). There is no consensus in the literature on an exact measurement, or even whom should do the measuring (Fletcher & Baldry, 2000). Other's have concluded that the construct of self-awareness simply cannot be accurately measured (Fleenor, et al., 1996). Two primary difficulties arise when considering a measurement of self-awareness: the rater and the standard of comparison.

The first problem lies in the question 'who should evaluate self-awareness?' Self-awareness is a construct of the self and by definition cannot be known by someone other than the self. Additionally, Wohlers and London (1989) have suggested that self-awareness is one of the most difficult competencies for others to evaluate. Self-ratings are also a problem because people cannot accurately evaluate their own self-awareness. A lack of one's self-awareness will result in the over or under estimation of his or her own level of self-awareness. Self-awareness is a necessary prerequisite for its accurate measurement, and therefore cannot be accurately measured by the self.

The second problem of measuring self-awareness is the standard of comparison. Researchers have argued that other's scores are not necessarily true scores, and should not be the standard of comparison for self-awareness (Alimo-Metcalfe, 1998). Murphy and Cleveland

(1995) argue that MSF should not be considered an instrument that gives a numerical value to employee performance, but rather a subjective measure of perceptions. MSF scores reflect a complex interaction between the goals of the rater (e.g. motivate leader with frank feedback, maintain interpersonal relationships), the context of the rating, the purpose of the evaluation, the performance level of the rater, and the relationship between rater and leader (Murphy, Cleveland, & Mohler, 2001; Saavedra & Kwun, 1993). Ratings should not be considered a reflection of an employee's true performance level, but rather a measure of people's *perception* of performance. Some competencies, such as communication, are completely dependent on the perception and interpretation of the rater. Thus, other's scores are mere measurements of perception and should not be the true measure of comparison for self-awareness.

Thus, the main difficulty in making the connection between self-awareness and leadership style is the conceptualization and measurement of self-awareness. Several in the literature have taken the theoretical leap of conceptualizing self-awareness as self-other agreement. In these instances the "standard" of measurement is the average score from all the raters (Atwater & Yammarino, 1992; Delmhorst, 2006). Leaders that are self-aware are better able to incorporate others perceptions into their self-evaluations.

Others have noted several flaws in conceptualizing self-awareness as self-other agreement (Fleenor, et al., 1996). Van Velsor et al. (1993) stated that the relationship between self-other agreement and self-awareness remains unresolved in the literature. Wohlers and London (1989) conducted a study measuring self-other agreement as well as a direct measure of self-awareness. Results indicated an inconsistent relationship between self-other agreement and self-awareness. Leaders that had the most inflated self-ratings (over-estimators) rated themselves highest on the direct measure of self-awareness, but received the lowest self-awareness ratings

by others. Additionally, it was the under-estimators that received the highest self-awareness ratings by others. In the current study there were no significant correlations between both measures of self-other agreement and the direct measure of self-awareness. This suggests that there are inconsistencies in measuring self-awareness, and that self-other agreement does not easily translate into self-awareness.

The current study does not make the theoretical leap of conceptualizing self-other agreement as a measure of self-awareness. Additionally, the author cannot assume the questionnaire used in the current study fully captures the construct of self-awareness. In fact the questionnaire in the current study may be subject to common-source bias (Dionne, Yammarino, Atwater, & James, 2002), where the same participants rated the independent measure of self-awareness and the dependent measure of transformational leadership. However, the current study did attempt to measure several variations of measurement in order to avoid common method variance (Ross & Gray, 2006). Little conclusions can be drawn from this study regarding the measurement of self-awareness, and it seems the questions surrounding its measurement remain unclear.

Predicting Transformational Leadership

Although *self-other agreement* D^2 and the direct measure of self-awareness significantly predicted transformational leadership, the strongest predictor of transformational leadership was the overall scores from the MSF survey. Almost 60% of the variance accounted for in transformational leadership was explained by the overall scores in MSF (see Table 11). There are two primary explanations for this finding: the halo effect and the overlap in leadership dimensions.

The halo effect is the failure of raters to discriminate between distinct and independent aspects of a ratee's behavior (Frone, Adams, Rice, & Instone-Noonan, 1986). Thorndike (1920,

p. 25) described the halo effect as the "marked tendency to think of the person in general as rather good or rather inferior and to color the judgments of the qualities by this general feeling." These early theorist would classify the halo effect as rater error. However, other studies have found support that the halo effect may be more than just rater error (Murphy, Jako, & Anhalt, 1993; Nathan & Tippins, 1990; Solomonson & Lance, 1997) .

Nathan and Tippins (1990) compared MSF ratings with the target leader's test scores on several abilities. These abilities overlapped with the competency dimensions on the MSF survey. Results from this study contradicted the traditional view that the halo effect is simply rater error. The halo effect did not appear to be contaminating the leader's performance ratings, rather the halo effect increased the predictability of the leader's performance. The authors concluded that halo effect did not turn out to be an error in predicting leadership performance, but instead there was considerable overlap in competency dimensions. An additional finding was that overall ratings of performance turned out to be a better predictor than the more specific dimension ratings.

Additional studies have supported the theory that there is an important conceptual distinction between true halo and halo error (Solomonson & Lance, 1997). True halo is the correlation among the actual behaviors being rated, which is due to the overlap and relatedness of the actual dimensions. For example, is it assumed that an effective leader is also a good communicator; therefore, ratings on the dimensions of leadership and communication should be correlated. Halo error inflates the correlations among dimensions beyond the level of true halo that may be present, due to the rater's impressions that the leader is generally good or generally inferior (Murphy, et al., 1993).

The halo effect may be contributing to the tendency for raters in the current study to rate similarly across competency dimensions. The high correlation between the MSF survey and the dimensions of leadership could lead one to assume that raters had a tendency to rate their leaders as generally good or generally inferior, regardless of the competency dimension. Although there may be halo error in raters' responses, there also seems to be some evidence of true halo due to the overlap in competency dimensions and leadership.

The second explanation for the overall MSF scores best predicting leadership is the overlap in leadership dimensions. Leadership may be so comprehensive that it encompasses most of the competencies in MSF surveys. Bean et al. (2006) conducted a study to look at the relatedness among MSF competencies. Of the 281 variations of competencies, there were four general clusters of competencies. These clusters could also be reduced to one general factor of leadership when performing a factor analysis, due to the high intercorrelations among competencies. These results support the theory that leadership encompasses several interrelated competencies. Thus, the results from the current study could be due to the interrelatedness of the MSF competencies and leadership, and less likely to be a result of halo error.

Gender Differences

Results from the current study indicated that female leaders were perceived as more transformational than the male leaders. Female leaders received significantly higher ratings on overall transformational leadership, idealized influence, individualized consideration and contingent reward. These findings are consistent with previous research that has found that females tend to be seen as more transformational leaders than their male counter parts (Kark, 2004; Trinidad & Normore, 2005). These findings hold true cross-culturally (Carless, 1998) and across industry (Druskat, 1994). It is also interesting to note that female leaders have a tendency

to rate their performance lower than males and be more receptive to feedback than males (Atwater & Waldman, 1998; Atwater & Yammarino, 1993).

Limitations and Future Research

There are some obvious limitations of the sample for this study. Due to the complexities of MSF, only 31 leaders were able to be recruited for the current study. Although there was a small sample size, several significant finders were evident. It would also be assumed that a larger sample would not change the null results of hypothesis two and three due to the minute effects. Another limitation is the homogeneity of the companies. Although from different industries, three of the four companies were from the same geographic location. Future research could look into the differences that may exist between industry and location.

The reported findings of gender differences may have been confounded by the disparate number of female leaders. Only 6 of the 31 leaders in the study were female. However, the finders were consistent with previous research (Atwater & Waldman, 1998; Carless, 1998; Druskat, 1994; Kark, 2004). There seems to be within the theory of transformational leadership a leniency toward feminine traits, despite the fact that some studies have found inconclusive results on gender (Van Engen, Van der Leeden, & Willemsen, 2001).

It is appropriate to reiterate the difficulties with the construct of self-other agreement and self-awareness. There is no consensus in the literature about the best way to measure self-other agreement, with currently over six methods demonstrated in the literature (Bracken, Timmreck, & Church, 2001). Studies have shown that some of the measurements of self-other agreement seem to be measuring different constructs (Fletcher & Bailey, 2003; Fletcher & Baldry, 2000). With the lack of consensus of how to measure self-other agreement and the difficulties in measuring self-awareness, it seems inappropriate to conclude a sure connection between self-

other agreement and self-awareness (Fleenor, et al., 1996; Fletcher & Bailey, 2003). The current study did not seek to resolve the theoretical issues connected with these concepts.

The current study was limited to a survey design, without experimental manipulations or a comparison group of leaders. Both the independent variables and dependent variables of this research were measured by a questionnaire survey. Future research could measure the dependent variable through a different method, such as measuring transformational leadership through observation, and thus avoiding common-method variance (Leithwood & Jantzi, 2006; Ross & Gray, 2006).

When using a survey design some of the measurements of the independent variable may become confounded with the dependent variable, due to common-source bias (Dionne, et al., 2002). Common-source bias is when participants rate both the independent variable and dependent variable similarly causing the two variables to be correlated with one another. This bias could possibly explain the effects of the direct measure of self-awareness and the overall MSF scores on transformational leadership. However, the common-source bias could not affect the results of the hypotheses because the measurements of agreement are calculated by a relationship between the raters' scores and leaders' score, not the raters' scores directly. Future research could test the additional findings from the current study using a split-sample design, which requires a larger sample of participants.

The greatest contributions of future research would be to further refine the construct of self-other agreement and self-awareness. Future research in the area of MSF and leadership development cannot go much further until there is an agreed measurement of self-awareness. An agreed upon operationalization of self-awareness is the first necessary step to linking leadership and self-awareness. Additionally, there needs to be an agreed and practical measurement of self-

other agreement before this area of research can make further contribution to the literature and assist companies in making sense of the discrepancies that exist between the leader's self-ratings and other raters.

REFERENCES

Alimo-Metcalfe, B. (1998). 360 degree feedback and leadership development. *International Journal of Selection and Assessment, 6*(1), 35-44.

Atkins, P., & Wood, R. E. (2002). Self- versus others' ratings as predictors of assessment center ratings: Validation evidence for 360-degree feedback programs. *Personnel Psychology, 55*(4), 871-904.

Atwater, L. E., & Brett, J. F. (2006). 360-degree feedback to leaders: Does it relate to changes in employee attitudes? *Group & Organization Management, 31*(5), 578-600.

Atwater, L. E., & Waldman, D. (1998). 360-degree feedback and leadership development. *Leadership Quarterly, 9*(4), 423.

Atwater, L. E., & Yammarino, F. J. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology, 45*(1), 141-164.

Atwater, L. E., & Yammarino, F. J. (1993). Personal attributes as predictors of superiors' and subordinates' perceptions of military academy leadership. *Human Relations, 46*(5), 645-668.

Avolio, B. J., & Bass, B. M. (2004). *Multifactor Leadership Questionnaire: Manual and sampler set (3rd ed.)*. Menlo Park, CA: Mind Garden.

Avolio, B. J., Waldman, D. A., & Yammarino, F. J. (1991). Leading in the 1990's: The four I's of transformational leadership. *Journal of European Industrial Training, 15*(4), 9-16.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147.

Baril, G. L., Ayman, R., & Palmiter, D. J. (1994). Measuring leader behavior: Moderators of discrepant self and subordinate descriptions. *Journal of Applied Social Psychology, 24*(1), 82-94.

Bass, B. M. (1985). Leadership: Good, better, best. *Organizational Dynamics, 13*(3), 26-40.

Bass, B. M. (1990). From transactional to transformational leadership: Learning to share the vision. *Organizational Dynamics, 18*(3), 19-31.

Bass, B. M. (1998). *Transformational Leadership: Industrial, Military, and Educational Impact*. Mahwah: Lawrence Erlbaum Associates.

Bass, B. M., & Avolio, B. J. (1994). *Improving Organizational Effectiveness through Transformational Leadership*. Thousand Oaks: Sage Publications.

Bass, B. M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology: An International Review, 40*(4), 437-454.

Bean, N., Williams, K., Jackman, B., Hardman, R., Bubb, R., Otto, S., et al. (2006). *Finding Underlying Competencies in Multi-Rater Assessments: A Multi-Dimensional Scaling Analysis Approach to Designing Feedback Instruments.* Paper presented at the Fulton Mentored Resrach Conference, Provo, UT.

Bernieri, F. J., Zuckerman, M., Koestner, R., & Rosenthal, R. (1994). Measuring person perception accuracy: Another look at self-other agreement. *Personality and Social Psychology Bulletin, 20*(4), 367-378.

Bono, J. E., Foldes, H. J., Vinson, G., & Muros, J. P. (2007). Workplace emotions: The role of supervision and leadership. *Journal of Applied Psychology, 92*(5), 1357-1367.

Borman, W. C. (1997). 360° ratings: An analysis of assumptions and a research agenda for evaluating their validity. *Human Resource Management Review, 7*(3), 299-315.

Bracken, D. W., Timmreck, T. W., & Church, A. H. (2001). *The Handbook of Multisource Feedback*. San Francisco: Josey-Bass.

Burns, J. M. (1979). *Leadership*. New York: Harper & Row.

Carless, S. A. (1998). Gender differences in transformational leadership: An examination of superior, leader, and subordinate perspectives. *Sex Roles, 39*(11), 887-902.

Chambers, E. G., Foulon, M., Hansfield-Jones, H., Hankin, S. M., & Micheals, E. G. (1998). The war for talent. *The McKinsey Quarterly, 3*, 44-57.

Cline, V. B., & Richards, J. M., Jr. (1960). Accuracy of interpersonal perception: A general trait? *The Journal of Abnormal and Social Psychology, 60*(1), 1-7.

Cohen, J., & Kazdin, A. E. (2003) *Methodological Issues and Strategies in Clinical Research (3rd ed.)*. Washington, DC American Psychological Association.

Conway, J. M. (1996). Additional construct validity evidence for the task/contextual performance distinction. *Human Performance, 9*(4), 309-329.

Cronbach, L. J., & Furby, L. (1970). How we should measure 'change': Or should we? *Psychological Bulletin, 74*(1), 68-80.

Dauenheimer, D., Stahlberg, D., & Petersen, L.-E. (1999). Self-discrepancy and elaboration of self-conceptions as factors influencing reactions to feedback. *European Journal of Social Psychology, 29*(5/6), 725-739.

Delmhorst, F. P. (2006). *Self-awareness and leadership development: Examining the impact of self-subordinate agreement on perceptions of leadership.* ProQuest Information & Learning, US.

Dionne, S. D., Yammarino, F. J., Atwater, L. E., & James, L. R. (2002). Neutralizing substitutes for leadership theory: Leadership effects and common-source bias. *Journal of Applied Psychology, 87*(3), 454-464.

Druskat, V. U. (1994). Gender and leadership style: Transformational and transactional leadership in the Roman Catholic Church. *Leadership Quarterly, 5*(2), 99-119.

Duval, S., & Wicklund, R. A. (1972). *A Theory of Objective Self Awareness*. New York: Academic Press.

Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes, 64*(3), 307-324.

Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal, 36*(6), 1577-1613.

Farh, J., & Dobbins, G. H. (1989). Effects of self-esteem on leniency bias in self-reports of performance: A structural equation model analysis. *Personnel Psychology, 42*(4), 835-850.

Feinberg, B. J., Ostroff, C., & Burke, W. W. (2005). The role of within-group agreement in understanding transformational leadership. *Journal of Occupational and Organizational Psychology, 78*(3), 471-488.

Fleenor, J. W., McCauley, C. D., & Brutus, S. p. (1996). Self-other rating agreement and leader effectiveness. *Leadership Quarterly, 7*(4), 487-506.

Fletcher, C., & Bailey, C. (2003). Assessing self-awareness: Some issues and methods. *Journal of Managerial Psychology, 18*(5), 395.

Fletcher, C., & Baldry, C. (2000). A study of individual differences and self-awareness in the context of multi-source feedback. *Journal of Occupational & Organizational Psychology, 73*(3), 303-319.

Foster, C. A., & Law, M. R. F. (2006). How Many Perspectives Provide a Compass? Differentiating 360-Degree and Multi-Source Feedback. *International Journal of Selection and Assessment, 14*(3), 288-291.

Frone, M. R., Adams, J., Rice, R. W., & Instone-Noonan, D. (1986). Halo error: A field study comparison of self- and subordinate evaluations of leadership process and leader effectiveness. *Personality and Social Psychology Bulletin, 12*(4), 454-461.

Furnham, A., & Stringfield, P. (1998). Congruence in job-performance ratings: A study of 360 feedback examining self, manager, peers, and consultant ratings. *Human Relations, 51*(4), 517-530.

Geyer, A. L. J., & Steyrer, J. M. (1998). Transformational leadership and objective performance in banks. *Applied Psychology: An International Review, 47*(3), 397-420.

Gibbons, F. X. (1983). Self-attention and self-report: The 'veridicality' hypothesis. *Journal of Personality, 51*(3), 517-542.

Gretchko, S. (2007). The growing talent crisis: Challenges and solutions. *Aon Consulting,* 1-10.

Guest, C. W., & Blucher, S. R. (1998). *Acuman Workstyles 6.0: Technical Report on Methods and Validity*. San Rafael, CA: Acumen Inc.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology, 41*(1), 43-62.

Hater, J. J., & Bass, B. M. (1988). Superiors' evaluations and subordinates' perceptions of transformational and transactional leadership. *Journal of Applied Psychology, 73*(4), 695-702.

James, L. R., Demaree, R. G., & Wolf, G. (1993). r-wg: An assessment of within-group interrater agreement. *Journal of Applied Psychology, 78*(2), 306-309.

Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology, 89*(5), 755-768.

Kark, R. (2004). The transformational leader: Who is (s)he? A feminist perspective. *Journal of Organizational Change Management, 17*(2), 160-176.

Katz, J., Arias, I., & Beach, R. H. (2000). Psychological abuse, self-esteem, and women's dating relationship outcomes: A comparison of the self-verification and self-enhancement perspectives. *Psychology of Women Quarterly, 24*(4), 349-357.

Leithwood, K., & Jantzi, D. (2006). Transformational school leadership for large-scale reform: Effects on students, teachers, and their classroom practices. *School Effectiveness and School Improvement, 17*(2), 201-227.

Liao, H., & Chuang, A. (2007). Transforming service employees and climate: A multilevel, multisource examination of transformational leadership in building long-term service relationships. *Journal of Applied Psychology, 92*(4), 1006-1019.

London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. *Personnel Psychology, 48*(4), 803-839.

London, M., & Wohlers, A. J. (1991). Agreement between subordinate and self-ratings in upward feedback. *Personnel Psychology, 44*(2), 375-390.

Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*(3), 280-296.

Morgeson, F. P., Mumford, T. V., & Campion, M. A. (2005). Coming full circle: Using research and practice to address 27 questions about 360-degree feedback programs. *Consulting Psychology Journal: Practice and Research, 57*(3), 196-209.

Moser, K., Schuler, H., & Funke, U. (1999). The moderating effect of raters' opportunities to observe ratees' job performance on the validity of an assessment centre. *International Journal of Selection and Assessment, 7*(3), 133-141.

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*. Thousand Oaks, CA US: Sage Publications, Inc.

Murphy, K. R., Cleveland, J. N., & Mohler, C. J. (2001). Reliability, Validity, and Meaningfulness of Multisource Ratings. In D. W. Bracken, Timmreck, T. W., & Church, A. H. (Ed.), *The Handbook of Multisource Feedback*. San Francisco: Jossey-Bass.

Murphy, K. R., Jako, R. A., & Anhalt, R. L. (1993). Nature and consequences of halo error: A critical analysis. *Journal of Applied Psychology, 78*(2), 218-225.

Nathan, B. R., & Tippins, N. (1990). The consequences of halo 'error' in performance ratings: A field study of the moderating effect of halo on test validation results. *Journal of Applied Psychology, 75*(3), 290-296.

Ross, J. A., & Gray, P. (2006). School leadership and student achievement: The mediating effects of teacher beliefs. *Canadian Journal of Education, 29*(3), 798-822.

Rothstein, H. R. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology, 75*(3), 322-327.

Saavedra, R., & Kwun, S. K. (1993). Peer evaluation in self-managing work groups. *Journal of Applied Psychology*, *78*(3), 450-462.

Sala, F. (2003). Executive blind spots: Discrepancies between self- and other-ratings. *Consulting Psychology Journal: Practice and Research, 55*(4), 222-229.

Sala, F., & Dwight, S. A. (2002). Predicting executive performance with multirater surveys: Whom you ask makes a difference. *Consulting Psychology Journal: Practice and Research, 54*(3), 166-172.

Sinha, N. (2004). *Self-other rating discrepancy and personality antecedents.* ProQuest Information & Learning, US.

Smircich, L., & Chesser, R. J. (1981). Superiors' and subordinates' perceptions of performance: Beyond disagreement. *Academy of Management Journal, 24*(1), 198-205.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*(4), 526-537.

Solomonson, A. L., & Lance, C. E. (1997). Examination of the relationship between true halo and halo error in performance ratings. *Journal of Applied Psychology, 82*(5), 665-674.

Sosik, J. J. (2001). Self-other agreement on charismatic leadership: Relationships with work attitudes and managerial performance. *Group & Organization Management, 26*(4), 484-511.

Sundvik, L., & Lindeman, M. (1998). Acquaintanceship and the discrepancy between supervisor and self-assessments. *Journal of Social Behavior & Personality*, *13*(1), 117-126.

Swann, W. B., Griffin, J. J., Predmore, S. C., & Gaines, B. (1987). The cognitive-affective crossfire: When self-consistency confronts self-enhancement. *Journal of Personality and Social Psychology, 52*(5), 881-889.

Swann, W. B., Pelham, B. W., & Krull, D. S. (1989). Agreeable fancy or disagreeable truth? Reconciling self-enhancement and self-verification. *Journal of Personality and Social Psychology, 57*(5), 782-791.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*(1), 25-29.

Trinidad, C., & Normore, A. H. (2005). Leadership and gender: A dangerous liaison? *Leadership & Organization Development Journal, 26*(7), 574-590.

Van Engen, M. L., Van der Leeden, R., & Willemsen, T. M. (2001). Gender, context and leadership styles: A field study. *Journal of Occupational and Organizational Psychology, 74*(5), 581-598.

Van Velsor, E., Taylor, S., & Leslie, J. B. (1993). An examination of the relationships among self-perception accuracy, self-awareness, gender, and leader effectiveness. *Human Resource Management, 32*(2/3), 249-263.

Warr, P., & Bourne, A. (1999). Factors influencing two types of congruence in multirater judgments. *Human Performance, 12*(3), 183-210.

Wohlers, A. J., & London, M. (1989). Ratings of managerial characteristics: Evaluation difficulty, co-worker agreement, and self-awareness. *Personnel Psychology, 42*(2), 235-261.

Yammarino, F. J., & Atwater, L. E. (1993). Understanding self-perception accuracy: Implications for Human Resource Management. *Human Resource Management, 32*(2/3).

Yammarino, F. J., & Atwater, L. E. (1997). Do managers see themselves as others see them? Implications of self-other rating agreement for human resources management. *Organizational Dynamics, 25*(4), 35-44.

APPENDICES

Appendix A: Background Questionnaire

Please provide the following information about the feedback recipient.

Please indicate the manager you are providing feedback for: _____

This person is my:

1 - Manager (Supervisor)
2 - Co-worker (Peer)
3 - Direct Report (Subordinate)

How long have you KNOWN this person? _____

How long have you WORKED with this person? _____

How well have you been able to observe this person's behavior at work?

1 - I do not have contact with this person in any setting
2 - I have limited contact with this person and it is in the same setting
3 - I interact with this person occasionally, usually in the same setting
4 - I interact with this person almost every day in more than one setting
5 - I interact daily with this person in a variety of settings

---

What is 360 Degree Feedback?

360 Degree Feedback is a confidential process where a target manager receives anonymous competency evaluations from a number of viewpoints in an organization (peers, direct reports, supervisors, etc.). The results help determine the participant's priorities for development.

Please be honest
In order to be most helpful, please answer the questions that follow in a completely candid and frank manner, without being unduly critical or uncritical.

The competencies you will be evaluating are:
Communication
Decision Making
People Management
Customer Focus
Goal Setting
Self-Awareness
Leadership Style

Appendix B: Multi-Source Feedback Survey

This contains the competencies and sample items companies choose from to develop their own MSF survey. The surveys contained a 5-point scale ranging from (1) *never* to (5) *always*, with the option of *no opportunity to assess*.

Please indicate the extent to which you agree or disagree that the statements below applies to the person you are rating. The person I am rating…

Communication
Listens carefully without interrupting
Solicits ideas, suggestions, and opinions from others
Creates a comfortable climate for airing concerns

Decision Making
Makes firm, resolute decisions
Is willing to take bold, calculated risks

Delegation
Delegates to employee effectively
Establishes a manageable workload for others

Innovation
Views obstacles as opportunities for change
Is attentive to new ideas
Inspires innovation in the organization

Team Orientation
Empowers others to find solutions to problems
Involves others in the decision making process

People Management
Acts fairly when dealing with employees
Will confront and address poor employee behavior

Task Management
Quickly gets to the essence of problems
Sets deadlines and expectations when assigning tasks
Produces quality work even under pressure

Personal Values
Keeps commitments
Shows commitment to the organization

Results Oriented
Is assertive in managing problems
Takes the initiative to get things done

Development
Maintains a timely schedule for reviews and follow-ups
Gives recognition to producers of high quality work

Customer Focus
Understands the needs of current and potential customers
Exceeds customers' expectations of agreed service

Resourcefulness
Makes good decisions when faced with incomplete information
Engages in flexible problem solving behavior

Valuing Diversity
Enjoys being with people different from themselves
Is sensitive to national and cultural differences
Is good at adapting business practices to other cultures

Goal Setting
Accomplishes long-term objectives by planning and taking the necessary steps
Sets clear performance standards

Technical Knowledge
Masters new technical knowledge necessary for the job in a timely manner
Has a sound understanding of the industry

Life balance
Balances work and personal life so that neither is neglected
Sets priorities in their private and professional life