

Winter 2013

# Exploring the structural relationships between personality and 360-degree feedback

Cole N. Napper

Follow this and additional works at: <https://digitalcommons.latech.edu/dissertations>

 Part of the [Personality and Social Contexts Commons](#), and the [Quantitative Psychology Commons](#)

---

**EXPLORING THE STRUCTURAL RELATIONSHIPS  
BETWEEN PERSONALITY AND 360-DEGREE  
FEEDBACK**

by

Cole N. Napper, B.A., M.A.

A Dissertation Presented in Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

COLLEGE OF EDUCATION  
LOUISIANA TECH UNIVERSITY

March 2013

UMI Number: 3570082

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

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

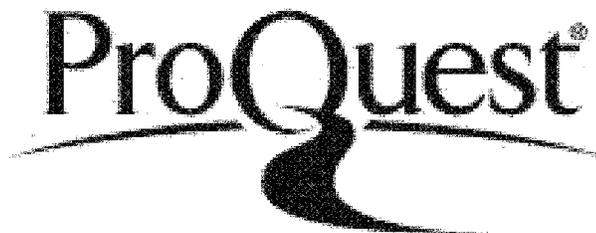


UMI 3570082

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

LOUISIANA TECH UNIVERSITY  
THE GRADUATE SCHOOL

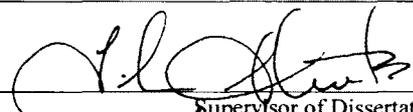
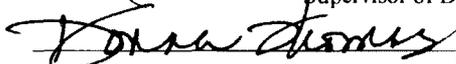
November 3, 2012

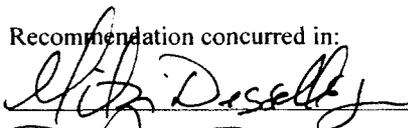
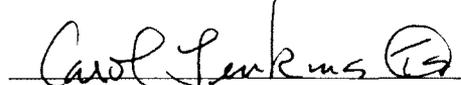
Date

We hereby recommend that the dissertation prepared under our supervision  
by Cole N. Napper

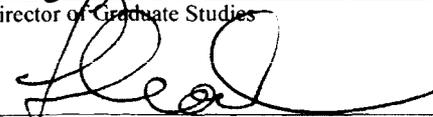
entitled Exploring the Structural Relationships between Personality  
and 360 Degree Feedback

be accepted in partial fulfillment of the requirements for the Degree of  
Doctor of Philosophy

  
\_\_\_\_\_  
Supervisor of Dissertation Research  
  
\_\_\_\_\_  
Head of Department  
Dept. of Psychology & Behavioral Sciences  
Department

Recommendation concurred in:  
  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_

Advisory Committee

Approved:   
\_\_\_\_\_  
Director of Graduate Studies  
  
\_\_\_\_\_  
Dean of the College

Approved:   
\_\_\_\_\_  
Dean of the Graduate School

## ABSTRACT

The process of using multiple sources or raters (i.e., self, supervisor, peers, subordinates, and others) in the assessment of managerial performance has been used pervasively in organizations with the primary goal of motivating behavioral change through feedback (Bracken, Timmrick, & Church, 2001). Multi-source or 360-degree feedback programs are especially suited to help measure behaviors related to performance and assess outcomes, such as leadership, interpersonal relationships, coaching, and communication (London & Smither, 1995). Typically, 360-degree feedback dimensions are measured by meta-categories of behavior called *competencies*. Bartram (2005) stated that these competencies could be defined as the search for characteristics that separate the best workers from the rest, usually related to characteristics, such as personality traits, that span across all jobs. However, DeNisi and Kluger (2000) stated that problems arise when managerial feedback is related to components of the *ideal self* (e.g., traits or individual differences) rather than feedback related to performance.

Spencer and Spencer (1993) proposed two kinds of competencies: skill-based competencies and trait-based competencies. Due to the negative outcomes associated with feedback disrupting the *ideal self* (i.e., decreased self-esteem, self-efficacy, and productivity), the present study sought to test whether 360-degree feedback competencies are related to personality traits of a person. Moreover, the present research tested Schmidt

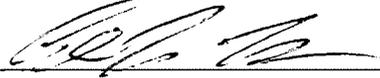
and Hunter's (1996) claim that interpreting the 360-degree feedback ratings as inter-rater correlations serving as reliability coefficients between ratings could represent a method of assessing the construct validity of 360-degree feedback ratings. Using confirmatory factor analysis (CFA), the present research modeled 360-degree feedback competencies by averaging across rater types (with and without self-ratings) and hierarchically across feedback items. Confirmatory models were then transformed into structural models in which personality characteristics of the *Big Five* were hypothesized to globally predict trait-based competencies, while not predicting skill-based competencies.

The present study indicates that hierarchical confirmatory models of the 360-degree feedback competencies have the most clear fit indices and validity coefficients. Mixed results were found for the hypothesis of personality characteristics of the Big Five predicting trait-based competencies, while the hypothesis regarding skill-based competencies was not supported. Detailed findings and implications of the research are discussed.

## APPROVAL FOR SCHOLARLY DISSEMINATION

The author grants to the Prescott Memorial Library of Louisiana Tech University the right to reproduce, by appropriate methods, upon request, any or all portions of this Dissertation. It was understood that "proper request" consists of the agreement, on the part of the requesting party, that said reproduction was for his personal use and that subsequent reproduction will not occur without written approval of the author of this Dissertation. Further, any portions of the Dissertation used in books, papers, and other works must be appropriately referenced to this Dissertation.

Finally, the author of this Dissertation reserves the right to publish freely, in the literature, at any time, any or all portions of this Dissertation.

Author 

Date 11-5-12

## **DEDICATION**

I would like to dedicate this dissertation to all those people who believed in me and those who did not. Without both, I would not be where I am today.

## TABLE OF CONTENTS

ABSTRACT.....	iii
DEDICATION.....	vi
LIST OF TABLES.....	x
LIST OF FIGURES .....	xiii
ACKNOWLEDGMENTS .....	xv
CHAPTER ONE INTRODUCTION.....	1
History of Performance Feedback .....	1
Self-Regulation Theory.....	5
Feedback Intervention Theory .....	6
History of 360-Degree Feedback .....	7
Relationship to Performance Appraisal .....	10
Rating Discrepancies in 360-Degree Feedback .....	12
Self-Appraisal Discrepancies.....	15
Criticisms of 360-Degree Feedback.....	20
Developmental Versus Decision-Making Feedback.....	21
Competency Models .....	23
Overview of a 360-Degree Feedback System.....	29
Current 360-Degree Feedback Research .....	34
Culture and 360-Degree Feedback.....	36

Positive Organizational Outcomes Using 360-Degree Feedback.....	37
Personality and 360-Degree Feedback.....	40
Five Factor Model.....	40
Other Personality Perspectives .....	42
Measures of Personality .....	44
Models of Personality and 360-Degree Feedback .....	46
Structural Equation Modeling.....	48
The Theoretical Basis of SEM.....	49
Construct Validity .....	53
Composite and Multiple Criterion .....	54
SEM, Personality, and 360-Degree Feedback .....	57
Relationship Between 360-Degree Feedback and Personality .....	59
Personality and Work Behavior.....	59
Theoretical Bases of 360-Degree Feedback .....	60
SEM and Competencies.....	62
Self-Rating Differences .....	63
Hypotheses.....	64
Hypothesis One.....	65
Hypothesis Two .....	65
Hypothesis Three .....	66
Hypothesis Four.....	67
Hypothesis Five .....	67
Hypothesis Six .....	68

CHAPTER TWO METHOD .....	70
Participants.....	70
Measures .....	70
GZTS/DFOS Personality Inventory.....	70
Relation to the Big Five .....	72
Strategic Success Model 360-Degree Feedback Assessment .....	72
Trait-Based Competencies.....	73
Skill-Based Competencies .....	74
Procedure .....	75
Data Screening .....	75
Data Analysis .....	75
CHAPTER THREE RESULTS .....	80
Confirmatory Factor Analyses.....	80
Business Acumen Higher-Order Construct .....	88
Drive for Results Higher-Order Construct.....	91
Managing Others Higher-Order Construct .....	94
Planning & Organizing Higher-Order Construct.....	98
Relationship Management Higher-Order Construct.....	102
Resilience Higher-Order Construct .....	105
Written Communication Higher-Order Construct .....	108
Structural Equation Models .....	114
Trait-Based Structural Models.....	114
Skill-Based Structural Models.....	120

CHAPTER FOUR DISCUSSION .....124

    Limitations .....134

    Suggestions for Future Research .....135

REFERENCES .....137

## LIST OF TABLES

Table 1	Relationship Between the GZTS/DFOS and the NEO-PI-R .....	45
Table 2	Comparative Model Fit Indices: Hypothesis One.....	81
Table 3	Item Content and Standardized Factor Loading Estimates: Hypothesis One .....	82
Table 4	Comparative Model Fit Indices: Hypothesis Two .....	83
Table 5	Item Content and Standardized Factor Loading Estimates Hypothesis Two .....	85
Table 6	Items and Standardized Factor Loading Estimates: Hypothesis Three .....	87
Table 7	Comparative Model Fit Indices: Hypothesis Four.....	88
Table 8	Item Content and Standardized Factor Loading Estimates: Business Acumen.....	89
Table 9	Discriminant Validity Squared Inter-Construct Correlations: Business Acumen.....	90
Table 10	Discriminant Validity Comparisons: Business Acumen .....	91
Table 11	Item Content and Standardized Factor Loading Estimates: Drive for Results .....	92
Table 12	Discriminant Validity Squared Inter-Construct Correlations: Drive for Results .....	93
Table 13	Discriminant Validity Comparisons: Driving for Results .....	94
Table 14	Item Content and Standardized Factor Loading Estimates: Managing Others.....	95
Table 15	Discriminant Validity Squared Inter-Construct Correlations: Managing Others.....	97

Table 16	Discriminant Validity Comparisons: Managing Others .....	97
Table 17	Item Content and Standardized Factor Loading Estimates: Planning & Organizing .....	99
Table 18	Discriminant Validity Squared Inter-Construct Correlations: Planning & Organizing .....	101
Table 19	Discriminant Validity Comparisons: Planning & Organizing .....	101
Table 20	Item Content and Standardized Factor Loading Estimates: Relationship Management .....	103
Table 21	Discriminant Validity Squared Inter-Construct Correlations: Relationship Management .....	104
Table 22	Discriminant Validity Comparisons: Relationship Management .....	105
Table 23	Item Content and Standardized Factor Loading Estimates: Resilience .....	106
Table 24	Discriminant Validity Squared Inter-Construct Correlations: Resilience .....	107
Table 25	Discriminant Validity Comparisons: Resilience.....	108
Table 26	Item Content and Standardized Factor Loading Estimates: Written Communication .....	109
Table 27	Discriminant Validity Squared Inter-Construct Correlations: Written Communication .....	110
Table 28	Discriminant Validity Comparisons: Written Communication .....	111
Table 29	Comparative Model Fit Indices: CFA Hypotheses Comparison .....	113
Table 30	Structural Model Results: Drive for Results.....	115
Table 31	Structural Model Results: Managing Others .....	116
Table 32	Structural Model Results: Planning & Organizing .....	117
Table 33	Structural Model Results: Relationship Management .....	119
Table 34	Structural Model Results: Resilience.....	120

Table 35	Structural Model Results: Business Acumen.....	121
Table 36	Structural Model Results: Written Communication .....	123

## LIST OF FIGURES

Figure 1	Comparison of Traditional Feedback Model to 360° Feedback Systems.....	11
Figure 2	A Diagram of Bracken’s (2001) Competency Model Outcomes .....	25
Figure 3	CFA Model for 360-Degree Feedback Competencies.....	65
Figure 4	CFA Model with No Self-Ratings. ....	66
Figure 5	CFA Model for GZTS/DFOS Personality Variables.....	66
Figure 6	SEM with Higher-Order Competency Factor.....	67
Figure 7	Personality Traits Predicting Trait-Based Competencies. ....	68
Figure 8	Personality Traits Not Predicting Skill-Based Competencies .....	69
Figure 9	CFA Model for Hypothesis One-Three Without Measured Variables. ....	77
Figure 10	SEM Model for Hypotheses Four-Six without Measured Variables .....	78

## ACKNOWLEDGMENTS

I would like to acknowledge all the professors and members of my committee who have assisted me in this dissertation. Special thanks to Dr. Tilman Sheets, Dr. Mitzi Desselles, Dr. Barry Babin, and Dr. Carol Jenkins. Without their assistance, this dissertation would not have been possible. Dr. Sheets has challenged me to be a better student from the very beginning, and for that I am grateful. Dr. Desselles has molded me into a better writer, editor, and researcher through many hours of editing and much effort. Dr. Babin has brought about a new-found passion for theoretical and empirical understanding of the world, and has tested me to defend my thought processes. And lastly, Dr. Jenkins has made this dissertation possible by providing me with the data and guidance to synthesize such large numbers of participants into meaningful results.

Also, a special thanks to all of my colleagues and friends in the I-O psychology doctoral program and fellow doctoral programs at Louisiana Tech University. Thank you for broadening my horizons, challenging me, calling me out on my mistakes, and having a couple of beers along the way. Each of you made the last couple of years in Ruston tolerable. My parents deserve a special thanks for putting up with me throughout graduate school, searching for jobs, and working on this dissertation. Mom and Dad, you are the best.

## CHAPTER ONE

### INTRODUCTION

Assessment, the foundation of 360-degree feedback, is considered an essential component in the delivery of organizational interventions (Fitzgibbons, 2003). Bracken, Timmreck, and Church (2001) pointed out that standardized psychological assessment is generally accepted to have begun around 1900 in Paris when physician, Alfred Binet, developed assessments for school children. Two techniques credited to Binet (i.e., standardized items and normed responses) still underlie most current methods of psychological assessment. From the 1920s through the 1950s, the development of psychological tests, surveys, inventories, and other instruments became a growing industry for psychologists (Bracken et al., 2001). An early example of the growth of the assessment business is Walter Dill Scott's man-to-man comparison scale (Paterson, 1922), which was used to assess employees on the job. Three decades later, Robert Bailey designed the first *multi-source feedback* survey (Bracken et al., 2001). The use of multiple sources in feedback eventually spread and was termed *360-degree feedback* (Bracken et al., 2001).

#### History of Performance Feedback

Although the origins of performance feedback may be lost to antiquity, the concept of performance feedback may be as old as work itself. However, the notion of

rating employees and giving quantified feedback is a more recent phenomenon. Bracken et al. (2001) note that almost a century ago, Thorndike observed that when supervisors rated their subordinates, the correlations among separate measures of performance were far too high. Thorndike postulated that inflated rating correlations may have been due to some form of error or bias. His observation spawned some of the first discussions about the appropriateness of using feedback ratings as a measure of job performance (Bracken et al., 2001).

The measurement of managerial performance has been a complex and ambiguous issue since the inception of modern assessment methods at the beginning of the twentieth century (Rainy, 1997). Many factors have contributed to the confusion in managerial assessment, including the sometimes conflicting roles that managers play, the number of constituents and stakeholders reporting to the manager, and the organizational, social, and political structures rooted in the work environment (Hassan & Rorhbaugh, 2009).

Ammon (1956) proposed two types of performance feedback for rating managers: (a) knowledge of results (KR) and (b) knowledge of performance (KP). He believed that both KR and KP lead to increases in learning and motivation. Ammon concluded that feedback consisting of both KR and KP should be provided to the employee from the supervisor. This supports the notion that supervisors should be the primary provider of performance feedback to their employees.

Hagan, Konopaske, Bernardin, and Tyler (2006) described traditional top-down assessment systems as consisting of one person, the direct supervisor, conducting a periodic evaluation of employee competence or performance over a specified period of time. However, when managers are rated by their supervisors, traditional performance

appraisals may have to be altered to be more effective, such as including self-ratings. The authors stated that because managers give and receive feedback, they have the optimal vantage point for observing and rating their own performance. Yammarino and Atwater (1993) pointed out that managers gain a more comprehensive perspective of their job performance when feedback is provided from different perspectives. DeNisi & Kluger (2000) stated that most managers believe that receiving feedback about job performance makes it more likely that their performance on the job will improve. However, certain types of feedback may be less effective, and others may even be harmful. The authors emphasized that feedback can also be harmful if it is not received well by the manager or if the manager perceives any type of bias in the feedback process. According to Nemeroff and Cosentino (1979) feedback recipients should have goals set for them by their superiors, but also be provided with the opportunity to set goals for themselves. The authors stated that this is because self-set goals lead to greater increases in performance than goals set by superiors while also increasing the perceived fairness of the feedback process.

Employees are interested in receiving feedback on their performance, and when they do not receive feedback from their supervisors, they will often seek feedback from other sources (Ashford & Cummings, 1983). Supervisors need to understand that feedback is an effective motivational technique and can lead to increased employee satisfaction (Hackman & Oldham, 1980). Performance feedback plays a role in organizational motivation and decision-making because it provides the opportunity for employees to make adjustments in their performance. This creates a reoccurring feedback loop in which individuals learn from the outcomes of their decisions or behaviors (DeNisi

& Kluger, 2000). In a study by Levy, Cawley, and Foti (1998), the authors found that participants generally preferred positive feedback. Similarly, participants tended to respond more favorably to feedback that is better than they anticipated. In concordance with the fundamental attribution error, Ross (1977) found that participants preferred being told that internal factors, such as personality traits, were responsible for their positive performance. Positive reactions were more likely to occur when individuals were given internal positive attributions compared to when positive attributions were more external. In contrast, the authors found that when performance ratings were not positive, participants did not want the attribution or responsibility for their performance to be internalized.

Internal and external attribution reactions to feedback are not the only weaknesses of traditional top-down feedback systems. Some researchers have argued that traditional appraisals are so dysfunctional that they need to be abolished (Coens & Jenkins, 2000). More than 70 percent of managers in one study admitted to intentionally giving inflated or deflated evaluations (Longnecker & Ludwig, 1990). Findings such as these call into question the validity of traditional performance feedback. Longnecker and Ludwig (1990) suggested that some of the problems with feedback systems are caused by organizational politics and the competition over the allocation of scarce resources within an organization.

Social psychologists have hypothesized that some of the flaws in performance appraisal may not be intentional. There is evidence that participants in performance appraisal create interactions that are mutually beneficial to all participants, and this can happen without the participants deliberately planning or explicitly communicating their

desire to do so (Nutt, 1993; Rabinowitz, Kelley, & Rosenblatt, 1966). If performance appraisal ratings are being affected by political motives, performance appraisals may have flaws that other methods of appraisal may avoid. Consequently, theories of why 360-degree feedback may be a reasonable alternative to traditional performance appraisal are discussed.

### **Self-Regulation Theory**

Higgins (1987) stated that within Self-Regulation Theory there are two aspects in which the self focuses its attention. Sometimes, we focus our attention on our *ideal self*, which is the self that we aspire to be, while at other times, we focus our attention on our *ought self*, which is the self that others expect us to be. DeNisi and Kluger (2000) hypothesized that the biggest problem with feedback interventions occurs when our attention is focused on the self-level rather than the task-level. In general, individuals prefer to work on tasks that are more likely to produce a focus on the ideal self. But, when working on tasks that individuals are *supposed* to do or are forced to do, individuals are more likely to focus on the *ought self*.

Focusing on the *ought self* typically leads an individual to concentrate on prevention of punishment and avoidance of pain and negative consequences. DeNisi and Kluger (2000) explained that when employees feel threatened by negative feedback they may seek to avoid punishment by improving performance (i.e., the *ought self* is under performing). Negative feedback that focuses our attention on the “ought self” is likely to improve performance because employees want to avoid punishment. The authors point out that when we receive positive feedback on the same *required* tasks, there is no subsequent improvement in performance, because there is no expected punishment to

avoid. Feedback interventions that focus our attention on the *ideal self*, however, can interfere with subsequent performance by diverting attention away from the task at hand and lead individuals to question their self-concept and their values. When feedback interventions are focused on the *ideal self*, an individual may internalize the negative feedback, which has been found to decrease performance, self-efficacy, and self-esteem (DeNisi & Kluger, 2000).

### **Feedback Intervention Theory**

Although performance feedback has generally been viewed as a useful tool for improving performance, in the literature, the results have been inconsistent. In their meta-analysis of performance feedback, Kluger and DeNisi (1996) found an overall modest positive effect of feedback on performance. One startling finding, however, is that 38 percent of performance feedback resulted in decreased performance. In response to these findings, Kluger and DeNisi (1996) proposed a model of performance feedback called Feedback Intervention Theory. Their theory is based on five assumptions: (a) behavior is regulated by comparison of feedback with a goal or standard, (b) goals or standards are arranged hierarchically, (c) attention is limited and only feedback that receives attention will regulate behavior, (d) attention is normally directed to a moderate level in the hierarchy, and (e) feedback interventions change the focus of attention to effecting behaviors.

According to DeNisi and Kluger (2000), the five assumptions outlined in Feedback Intervention Theory trigger three specific reactions in terms of performance feedback. The first reaction occurs when a person notices a gap between feedback and some goal, and most often a person will try to reduce the gap. The second reaction

depends on the level of the goal toward which feedback is directed. Feedback may be directed at the self-concept level, the task level, or the task-learning level. Similar to Self-Regulation Theory, the self-concept level may be confronted if feedback is provided inappropriately. In the task level, a person directs attention toward the task itself, and the recipient works to reduce the gap between actual and desired performance. In the task-learning level, the level of the goal that influences behavior depends on where the attention is focused. Typically, attention is focused on the task itself, but feedback interventions can direct attention to different levels, depending on the goals of the intervention (DeNisi & Kluger, 2000).

Consequently, due to the popularity of generalized feedback in the workplace, many feedback-related processes have emerged (Latham & Wexley, 1994). Feedback interventions, such as traditional top-down performance appraisal, have evolved to meet the ever changing landscape of organizational structures, such as the flattening of organizations (Pollitt, 2005) and the increased use of teams (Roberts, 1995). London and Smither (1995) stated that feedback that is gathered from multiple perspectives may have led to the rise in a performance feedback system called 360-degree feedback. Moreover, 360-degree feedback has its roots in a concept termed multi-source feedback (Latham & Wexley, 1994).

### **History of 360-Degree Feedback**

London and Smither (1995) hypothesized that multi-source feedback programs are especially suited to measure behaviors related to components of job performance, such as leadership, interpersonal relationships, coaching, communicating, and maintaining good working relationships. Briefly, multi-source feedback can be described

as involving the use of multiple sources (i.e., self, supervisor, peers, subordinates, and others) in the assessment of individuals with the primary goal of motivating behavior change through feedback (Bracken et al., 2001; Tornow, 1993). The concept of multi-source feedback is grounded in the philosophy and practice of survey feedback (Nadler, 1977) and performance appraisal (Latham & Wexley, 1994). By utilizing the strengths of both of these concepts, multi-source feedback attempts to leverage the unique perspectives of employees from different levels within an organization (and on occasion incorporating members external to the organization, such as customers) to provide diverse performance feedback (London & Beatty, 1993). Early proponents of 360-degree feedback systems suggested that 360-degree feedback be used primarily for developmental purposes. The goal of these systems is to enrich employees' experiences and identify employees' strengths and weaknesses from a range of perspectives. The 360-degree feedback process is expected to create leaders that are well-adjusted and ready to be promoted (London & Smither, 1995). One of the key structural strengths of using 360-degree feedback is that the data generated provides a more comprehensive picture of a manager's performance in contrast to the singular lens of traditional, top-down feedback systems (Fletcher, 1999). Multiple viewpoints generate a more balanced assessment of managerial performance.

Some authors have suggested that the interaction between self-assessments and other-ratings may echo findings similar to those of other theoretical perspectives discussed in the feedback literature, such as Prospect Theory (Kahneman & Tversky, 1979), Control Theory (Carver & Scheier, 1981), and Feedback Intervention Theory (Kluger & DeNisi, 1996). All of these approaches suggest that when individuals receive

negative feedback (defined as self-ratings higher than other-ratings), they will be motivated to reduce the discrepancy and make efforts to improve their performance (Bailey & Austin, 2006). Due to the disparity between self-ratings and the performance ratings from others, Bailey and Austin (2006) asserted that cognitive dissonance will motivate an individual to reduce the gaps in performance, which may lead to an increase in the individual's motivation to perform. However, 360-degree feedback has problems associated with it that are similar to those found in traditional feedback systems. For example, several aspects of 360-degree appraisals have often increased the likelihood that the focus of the feedback will move towards the ideal self, which Higgins (1987) hypothesized will challenge an individual's self-concept. DeNisi and Kluger (2000) also pointed out, that 360-degree systems may also encounter problems leading to recipients internalizing the feedback which may lead to productivity loss.

In many organizations, 360-degree appraisals are administered only once and never repeated (London & Smither, 1995). Conducting a 360-degree feedback intervention only once makes it difficult for employees to know whether their performance is improving over time. Conversely, receiving feedback on multiple occasions has been found to improve managerial performance over time (London & Smither, 1995). London and Beatty (1993) pointed out that organizations may be missing out on some of the advantages of using 360-degree feedback systems. The authors stated that in practice 360-degree feedback could often be more accurately described as *270-degree feedback*. Major data sources (e.g., customers, subordinates, etc.) are often excluded. Excluding feedback from customers leaves performance gaps that other raters may fail to take into account (London & Beatty, 1993). Even though 270-degree

feedback interventions may have their shortcomings, these feedback systems still add value to an organization's performance assessment system.

From an organizational perspective, 360-degree feedback systems should add value to organizations, and the benefits accruing from 360-degree feedback should be addressed in the needs assessment phase (London & Beatty, 1993). The term *valued-added* refers to an initiative's direct contribution to a firm's competitive advantage; whereas, competitive advantage is defined as providing a product or service perceived by its customers as contributing to the organization and market in a way that is unique and difficult for a competitor to readily duplicate (Ulrich & Lake, 1990). Potential ways for 360-degree feedback programs to add value to organizations include: (a) providing better customer-centric data, (b) developing high potential leaders, and (c) increasing overall job performance (London & Beatty, 1993). Also, 360-degree feedback can enhance communications between feedback recipients and stakeholders while serving as input for merit evaluation and compensation adjustment (Bernardin & Beatty, 1987; McEvoy & Buller, 1987). The introduction and repeated use of a multi-source feedback program can also redefine the way employees think about their performance and alter their schemas about leadership (London & Smither, 1995). Redefining employee schemas may have an effect on employees' views of the performance appraisal system.

### **Relationship to Performance Appraisal**

Supervisory ratings are often the sole source of evaluative data in traditional performance appraisal, and these ratings are used for making decisions, such as performance-based pay, contingent reward structures, opportunities for promotion, and other supervisory decisions (London & Beatty, 1993). While the use of performance

appraisals is almost ubiquitous, the incidence of 360-degree feedback programs is on the rise (Lin, 2012). Toegel and Conger (2003) stated that the increase in 360-degree feedback programs may be due to the greater relevance of leader and manager development programs and the flattening of organizations. In addition, the increasing use of 360-degree feedback programs may also be attributed to organizations doing a better job at recognizing the complexity of management and valuing input from different sources (Becker & Klimoski, 1989). When surveyed, 90 percent of human resource executives said that, if given the opportunity, they would modify, revise, or even eliminate the performance appraisal system currently used in their organizations and move toward a more multi-source approach (Toegel & Conger, 2003). Figure 1 compares the conceptualization of traditional performance ratings to 360-degree feedback ratings.

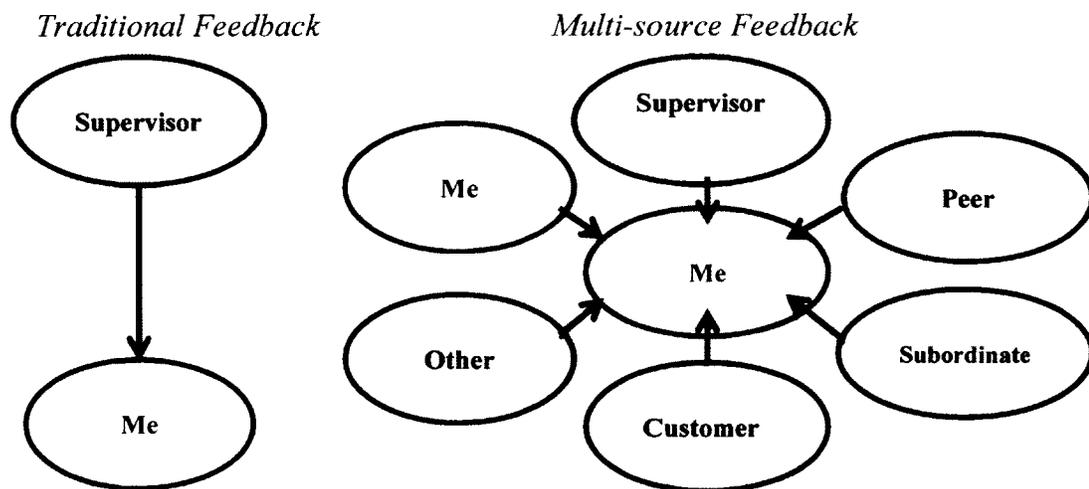


Figure 1 Comparison of Traditional Feedback Model to 360° Feedback Systems

Schippmann et al. (2000) estimated that between 75 and 80 percent of companies use some form of competency-based 360-degree feedback system. Using a competency-based approach, unique perspectives on leadership and management can be

tailored to fit management profiles within an organization, giving the organization a competitive advantage (Ulrich & Lake, 1990). The use of specialized leadership and management competencies has transformed the fundamental building blocks of workforce planning and succession management initiatives (Becker & Huselid, 1999). This transformation may explain why 360-degree feedback programs are becoming more popular developmental tools compared to traditional performance feedback measures. The use of multiple raters increases the power of 360-degree feedback, while still giving recipients a chance to express their own opinions about their performance. Self-ratings are especially suited to identify these situational constraints when compared to supervisor-ratings (Bernardin & Beatty, 1984). However, differences between self- and other-ratings can lead to decreased effectiveness of 360-degree feedback over time.

### **Rating Discrepancies in 360-Degree Feedback**

Van Velsor, Taylor, and Leslie (1993) suggested that in order for 360-degree feedback to be effective, feedback recipients must believe that the feedback is accurate, representative of the different rater views of performance, and build consensus among the rater groups. The authors pointed out that 360-degree feedback captures performance ratings from multiple perspectives, and different organizational stakeholders should be invested in the process for performance feedback to be effective. According to the authors, consensus cannot be built among the different rater groups if there is incongruence in the ratings. They defined rating incongruence as the degree to which ratings from multiple sources are dissimilar from each other.

Many opportunities exist for rater incongruence to occur between rater groups. Although rating incongruence in 360-degree feedback can exist for valid reasons

(Tornow, 1993), a high degree of incongruence between ratings is generally considered undesirable because it brings into question the validity of the ratings (Borman, 1997). Regardless of the flaws that may occur from rater incongruence, ratings from different perspectives appear to capture unique variance within performance (Borman, 1997). Also, including raters from different perspectives has been found to be a more valid assessment of performance than traditional, top-down performance appraisal ratings of managers (Mount, Judge, Scullen, Sytsma, & Hezlett, 1998). While the use of multiple raters can be seen as a strength of 360-degree feedback assessments, because of the unique variance each rater captures, multiple raters may also be barriers to the interpretation of feedback because of rater disagreements. If contradictory differences in ratings exist, it could be difficult for the feedback recipient to act on the feedback results.

Before the beginning of the twenty-first century, very few organizations used 360-degree feedback, and the supervisor was the traditional source of all feedback for employees (Toegel & Conger, 2003). Discrepancies in ratings between the supervisor and others did not exist because the supervisor was the only person who rated performance. With the introduction of 360-degree feedback systems, peers, subordinates, customers, and the person being rated all became equal participants in rating worker performance. Bernardin, Dahmus, and Redmon (1993) found that when ratings were collected from the supervisor, peers, and subordinates, feedback became more useful to recipients than if it had come solely from the supervisor. The researchers found that developmental feedback was also less likely to be ignored if it included peer ratings along with supervisor ratings. This is because peers and the supervisor have more power and status in the organization than do subordinates (Bernardin et al., 1993). More recently, Bailey and Fletcher (2002)

found that the incremental predictive value of supervisors and peers, over other sources of feedback, has previously been underestimated in the 360-degree feedback literature. Brett and Atwater (2001) concluded that feedback from subordinates does not appear to influence reactions as much as those from supervisors and peers. Their findings contradicted the literature supporting the reliability of subordinate ratings being one of the best predictors of feedback impact (Brutus, Fleenor, & McCauly, 1999) and reactions toward the feedback process (Maurer et al., 2002). These contradictory findings could lead one to believe that discrepancies and inconsistencies created by rater incongruence could have detrimental effects on the 360-degree feedback process and manager perceptions.

One theoretical reason for the effectiveness of 360-degree feedback appraisals is found in Social Comparison Theory (Festinger, 1954). Social Comparison Theory states that people believe that rewards are based on the differential possession of abilities and competencies. Due to most jobs lacking objective measures of performance, people tend to observe and compare themselves to the task-relevant abilities and competencies of others (Miller & Cardy, 2002). Managers with inconsistent patterns of feedback from their supervisor, peers, and subordinates may find the feedback confusing and unhelpful; therefore, rendering the feedback less effective as a tool of behavior change and managerial development (Miller & Cardy, 2002). Compared to a traditional performance appraisal, the employee receiving multi-source feedback has considerably more information to interpret and integrate than is usually provided in top-down appraisals. This amount of information can be confusing and de-motivating for managers if not presented effectively (London & Smither, 1995).

In their meta-analysis of 360-degree feedback ratings, Conway and Huffcutt (1997) found that managers who had been rated by their subordinates, supervisor, peers, and themselves had between-source correlations as high as .80 for supervisor and peer ratings, and no lower than .57 for subordinate-supervisor ratings. Although levels of inter-rater agreement only have moderate-to-low levels of reliability, perfect agreement is not necessary for a participant to gain insight from raters. Certain levels of rater disagreement may actually be seen as effective tools for motivating managers to improve isolated dimensions of performance (Conway & Huffcutt, 1997). For example, managers may receive feedback that their *relationship management* needs to be improved with their peers and subordinates, but is currently at acceptable levels with their supervisor. Although occasional rater disagreement can be seen as positive, certain rater discrepancies may be systemic and detrimental to the validity of multi-source ratings.

### **Self-Appraisal Discrepancies**

One of the most common discrepancies found in 360-degree ratings is the discrepancy between self-ratings and ratings by others (London & Smither, 1995). According to Thornton (1980), there are many reasons why self-appraisals may disagree with other appraisals, one of which is egocentric bias. Egocentric bias is described as when people claim more personal responsibility for the results of a joint action than an outside observer would credit them (Thornton, 1980). Two theories that underlie the use of self-appraisals are Bem's (1972) theory of self-perception and Bandura's (1986) social cognitive theory. Self-Perception Theory states that, just as we often infer other people's attitudes by observing their actions, we determine our own attitudes by observing our own actions (Bem, 1972). Likewise, Social Cognitive Theory states that while observing

our own behavior we may create self-set goals (Bandura, 1986). Once these goals are set, we use self-monitoring techniques to measure our progress towards the goals, administering self-set awards and punishments when necessary. Self-observation and self-monitoring may both been seen as increasing the salience of self-ratings in 360-degree feedback (London & Smither, 1995). Egocentric bias is seen in both theories because self-appraisals only focus on the personal contributions, goals, and outcomes of performance, and do not focus on others contributions to performance outcomes.

Within both Self-Perception Theory and Social-Cognitive Theory lays the concept of self-awareness. A person's self-awareness may explain the accuracy or inaccuracy of self-appraisals (London & Smither, 1995). One aspect of self-awareness that may affect 360-degree feedback self-ratings is how people see themselves in relationship to their peers. Festinger's (1954) Social Comparison Theory states that self-evaluations tend to entail absolute judgments of the self, and these judgments of comparison tend to be flawed. Self-appraisals that require relative judgments of ability compared to others tend to yield better approximations ability (Latham & Wexley, 1994). While using absolute judgments within self-appraisals, there appears to be some evidence that judgments improve with practice, especially practice that includes feedback on the accuracy of ratings (Latham & Wexley, 1994). Of all the different rating methods within 360-degree feedback, self-ratings tend to consistently be the most inflated (Jones, Rhodewalt, Berglas, & Skelton, 1981). Control Theory posits that self-rating inflation may not be a problem (Carver & Scheier, 1981). The authors state that when self-ratings are inflated compared to other-ratings, people may be more motivated to make changes in their

behavior. However, this does not change the possibility that self-appraisals may have been inaccurate all along.

Empirical research findings have consistently shown self-appraisals to be poor indicators of performance (London & Smither, 1995). A meta-analysis conducted by Conway and Huffcutt (1997) concluded that ratings from all sources except self-ratings have significant levels of inter-rater reliability. Their findings show an inter-rater correlation of .30 for subordinates, .37 inter-rater correlations for peers, and .50 inter-rater correlations for supervisors. Consequently, one could surmise that those with external viewpoints to the feedback recipient have moderately more consistent and valid performance ratings than the self-ratings (Conway & Huffcutt, 1997). Viswesvaran, Ones, and Schmidt (1996) found similar results with inter-rater correlations at .52 for supervisors and .42 for peers. In another meta-analysis conducted by Harris and Schaubroeck (1988), the authors found that the discrepancies between self- and other-ratings are consistent among all rater types. The meta-analysis showed non-significant correlations between self-supervisor ratings, self-peer ratings, and self-subordinate ratings.

This finding was similar to earlier conclusions from Mabe and West (1982) who found low correlations between self and other relationships, usually due to managers rating themselves higher than others rated them, decreasing the correlation. Outside of self-ratings, correlations of within-rater agreement in 360-degree feedback dimensions tend to range from about .30 to .50 (Conway & Huffcutt, 1997). Also, agreement between rater groups tends to be moderate as well, excluding self- and other-ratings (Harris & Schaubroeck, 1988). Agreement between-rater groups and within-rater groups has been

consistent, with the only inconsistency being self-ratings. Multi-source feedback ratings capture unique variance of managerial performance samples from all around the participant, providing a complete view of the participant's strengths and potential development areas (Nowack & Mashihi, 2012), but the validity of self-ratings continue to be a psychometric challenge.

Cheung (1999) noted potential problems can arise from conceptual disagreements that occur between raters. Raters may conceptualize performance in different ways. Rating effects are more strongly associated with individual raters rather than the rater's role (Mount et al., 1998). Using confirmatory factor analysis, Maurer, Raju, and Collins (1998) found that factor loadings are invariant across rater sources, meaning raters in different roles share a common conceptualization of performance dimensions. The most common index of between-rater agreement in the literature is the intercorrelation between two sets of values (e.g., the correlation of self-ratings and other-ratings). This form of agreement is often moderate, with an average finding around .20, but sometimes up to .30 (Warr & Bourne, 1998). Self and other rating differences have been found to be stable over time when rating different dimensions of competencies, such as skill-based or personality-based competencies (Nilsen & Campbell, 1993).

However, one of the issues of not including self-ratings in multi-source feedback is employee's buy-in. Farh, Werbel, and Bedeian's (1988) research indicated that when given the opportunity to evaluate their own performance, employees tend to be more satisfied with self-rating appraisal systems compared to appraisal systems with no self-ratings. Self-appraisals have also been shown to increase ratee participation in the appraisal interview (Bernardin & Beatty, 1984) and perceptions of procedural justice and

fairness (Farh et al., 1988). However, after receiving feedback from other raters, subsequent self-ratings seem to more closely resemble those of other raters (Atwater, Rouch, & Fischthal, 1995).

Advocates of 360-degree feedback assert that repeated interventions can decrease the gap between self- and other-ratings. Nilsen and Campbell (1993) stated that the discrepancies between self- and other-ratings are stable over time until performance feedback is administered. After feedback, gaps tend to decrease over time. Other solutions for minimizing disagreement between self- and other-ratings include: conducting a job analysis to remove ambiguity (Campbell & Lee, 1988), comparing appraisals against objective criteria (Lane & Herriot, 1990), and providing employees with a frame of reference for their ratings (Farh & Dobbins, 1989).

In terms of self-rating, some individuals have a bias for self-enhancement (Jones et al., 1981). London and Smither (1995) stated that this bias may be due to an inherent predisposition for people to inflate their self-concept and exaggerate their accomplishments or talents. Inflated self-evaluations can be a problem for individuals because they have been found to be related to career failure (McCall & Lombardo, 1983) and low performance (Yammarino & Atwater, 1993). Over-rating is not merely a function of self-esteem. Farh and Dobbins (1989) found that certain personality traits can lead to over-rating, such as those related to narcissism. Personality-behavior links are stronger when behavior is recorded through self-ratings rather than supervisor-ratings, and when self-ratings are more positive than supervisor-ratings significant relationships to personality traits can be found (Warr & Hoare, 2002). Consequently, it should be noted that rater incongruence is not the only shortcoming of 360-degree feedback mechanisms.

### **Criticisms of 360-Degree Feedback**

Although 360-degree feedback has been referred to as one of the most significant contributions to management practice of the last 20 years (Atwater & Waldman, 1998; Chappelow, 2004; London & Beatty, 1993; Richardson, 2010), this form of feedback does not come without criticism. Similar to the halo effect, Buda, Reilly, and Smither (1991) found that once a ratee has been categorized by raters, raters will be more likely to recall information about the ratee in a way that is consistent with their initial categorization of the ratee. Findings such as this suggest that raters might not notice or recall small to modest improvements in ratee performance (Seifert, Yukl, & McDonald, 2003). From a developmental feedback perspective, when a 360-degree feedback intervention is conducted on multiple occasions, unless there have been drastic improvements between Time One and Time Two, ratings are unlikely to change (Buda et al., 1991). If recipients of the feedback do not see changes in their ratings despite small to moderate changes in behavior, they may become disappointed and disenfranchised with the system.

Interestingly, Avery (2000) found that feedback recipients do not necessarily improve more on the dimensions on which they were rated the lowest and in most need of improvement. When individuals do not improve on their weakest areas, raters may overlook improvements in other areas and give poor performance ratings across all dimensions (e.g., *negative halo effect*). Hezlett and Ronnkvist (1996) stated that without the proper action-planning and feedback to raters, the probability of observed behavior change is very low. With the amount of time and money typically invested into multi-source feedback interventions, failure to see developmental improvement can be

very expensive for an organization. Emphasizing the needs assessment phase of the feedback intervention will encourage organizations to weigh the benefits of the system and make sure that it is rolled out effectively. Furthermore, even though criticisms of 360-degree feedback exist, there may be hope for these systems. In a meta-analysis conducted by Seifert et al. (2003), the authors found nearly all of the effect sizes for subordinates, peers, and supervisors show feedback rating improvements between Time One and Time Two ratings. Although the magnitude of improvement between ratings was moderate, this finding exhibits some evidence of the efficacy of implementing 360-degree feedback interventions for organizations.

### **Developmental Versus Decision-Making Feedback**

Evidence of the effectiveness of 360-degree feedback can be seen in the transition over the past few decades from using 360-degree feedback as a developmental tool to using it as a performance appraisal or decision-making tool (Bettenhausen & Fedor, 1997; Fletcher & Baldry, 2000; London & Smither, 1995; Waldman, Atwater, & Antonioni, 1998). Advocates of using 360-degree feedback as a development tool only (Waldman et al., 1998) see its use for decision-making as one of the reasons why improvement is not always the universal outcome of the feedback process.

Designed primarily as a system for management development, 360-degree feedback was not originally intended to be used as a decision-making tool for promotions, dismissals, or compensation (Bracken et al., 2001; Mount et al., 1998; Waldman et al., 1998). Concern has been raised about the damage that may have been inflicted by using 360-degree feedback for performance appraisal (Dalton, 1997; DeNisi & Kluger, 2000; Toegel & Conger, 2003). In regards to 360-degree feedback for manager

performance improvement, multi-source feedback may provide a more comprehensive picture of managerial performance in contrast to singular ratings of the supervisor in performance appraisal (Fletcher, 1999). One study found that 34 percent of subordinates indicated they would have rated their managers more accurately had the feedback been used for performance development and not performance appraisal (London, Wohlers, & Gallagher, 1990). The question still remains as to whether the ratings on 360-degree feedback measures can be interpreted as a more valid performance criterion compared to traditional performance appraisal systems.

Consensus within the literature discussing the use of 360-degree feedback endorses using the feedback for strictly developmental purposes (London & Beatty, 1993; London & Smither, 1995; Morgeson, Mumford, & Campion, 2005). Even with consensus supporting developmental uses, some authors argue that 360-degree feedback should be used in decision-making processes as well as development. Bracken (2006) states that even though most of the research literature advocates using 360-degree feedback for developmental purposes, perhaps the system's potential is not fully being utilized for organizational benefits. The author argued that using 360-degree feedback for decision-making (e.g., performance management or succession planning) could be beneficial for companies and managers. He hypothesized that decisions, such as promotion, executive selection, and performance-based incentives, could all be tied to some sort of aggregate rating created by 360-degree feedback.

Toegel and Conger (2003) argued for two distinct models of 360-degree feedback, one for management development and one for performance feedback. The management development tool could rely more heavily upon qualitative feedback and competency

development. The performance appraisal feedback tool could be designed around quantitative feedback and measuring performance outcomes. The researchers argued that 360-degree feedback could be used successfully for both purposes; only the structure of the feedback would change depending on the circumstances for which the feedback was being used. Conversely, critics stated that using 360-degree data for performance appraisal makes the developmental process potentially *punitive* and one that is *forcing* instead of *enabling* change (Pollman, 1997).

However, Bracken (2006) acknowledged that for 360-degree feedback to be successful (for developmental or decision-making purposes) certain critical factors are necessary: support from top management, validated competency models, rater training, rater accountability, rater anonymity, organization-wide implementation, easy-to-use feedback and reporting mechanisms, action-planning, and ratee accountability (Bracken, 2006). Whether using 360-degree feedback systems for development or decision-making, the competencies selected for measurement are important to the effectiveness of 360-degree feedback.

### **Competency Models**

Specifically aimed at developing employees, one trend in business and research is the concept of work-related competencies. Bracken et al. (2001) described competencies as making up an umbrella category that represents a combination of skills, knowledge, abilities, values, and other individual difference characteristics necessary for effective performance. A competency is a feature that refers to a form of human capital or human resources that can increase productivity (Beheshtifar & Moghadam, 2011), and individual differences that can lead to higher performance (Lustri, Miura, & Takahashi, 2007).

Competencies typically represent the behavioral expression and trait-oriented combination of many individual characteristics necessary for success. Competencies can be thought of as a search for characteristics that separate the best workers from the rest, and these characteristics typically describe traits necessary across all jobs while ignoring tasks (Bartram, 2005). Examples of content categories that make up competencies include: customer focus, results orientation, innovation, leadership, collaboration, change orientation, and communication (Reilly & McGourty, 1998).

The goal of competency models is to identify organizationally-valued personal characteristics required of individual employees by jobs or roles (Brannick, Levine, & Morgeson, 2007). Competencies are powerful strategic business tools because they can serve as a framework for relating employee success requirements to the overall competencies and capability of the organization (Bracken et al., 2001). Another strategic advantage of competencies is that they provide a common language for defining, communicating, and evaluating employee behavior (Reilly & McGourty, 1998). Various job analysis techniques are available for identifying competencies, which give practitioners the appropriate steps to build competency models relevant to their organization and employees (Reilly & McGourty, 1998). One example of how competency models are developed to drive organizational outcomes is outlined by Bracken et al. (2001). The researchers stated that competencies developed from knowledge, skills, abilities, values, and individual differences can link business strategy to multi-source feedback (MSF) and organizational outcomes.

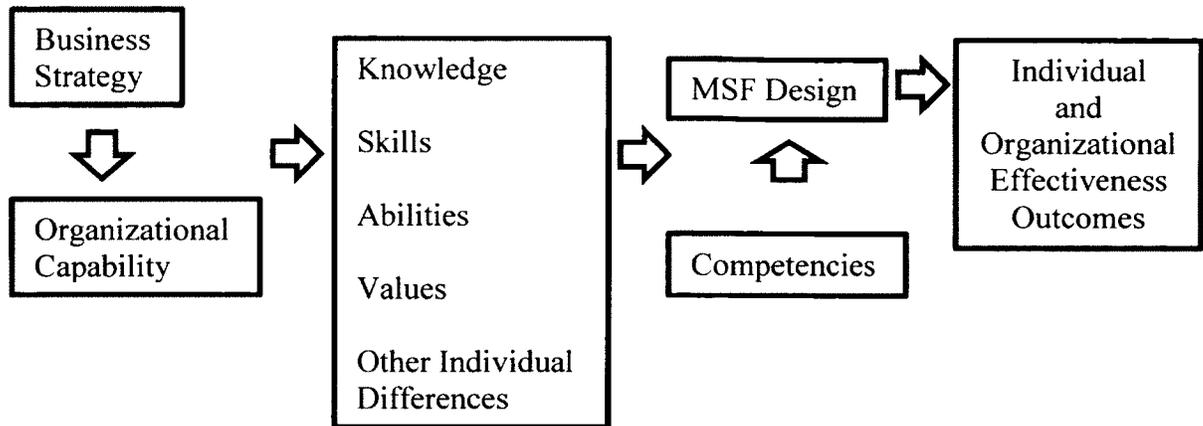


Figure 2 A Diagram of Bracken's (2001) Competency Model Outcomes

As previously stated, a competency can be defined as the underlying sets of skills, knowledge, personal characteristics, and abilities needed to effectively perform a role in the organization and meet organizational strategic objectives (Latham & Wexley, 1994). Brannick et al. (2007) stated that the notion of managerial or leadership success is at the core of competency modeling, and the topic has frequently been applied in the organizational research and consulting realm. In recent years, many companies have used the concept of competencies and competency models to define broad behavioral capabilities necessary to achieve organizational objectives.

A survey by Schippmann et al. (2000) found that 70-80 percent of companies were using some form of competency-related strategies for selection and development. Given the widespread use of competencies, there seems to be no universally accepted method or taxonomy of competencies (Bartram, 2005). However, the consulting firm SHL has attempted to fill the gap in the research regarding a universal taxonomy of competencies. Bartram (2005) stated that SHL has identified the *great eight* competency factors, with the eight competencies being composed of 20 sub-competencies and 112

sub-components. Using confirmatory factor analysis, Abdullah, Musa, and Ali (2011) attempted to develop another universal measure of competencies. The researchers found a 12-factor structure with three types of competencies based on previous research: behavioral competencies (Spencer & Spencer, 1993), technical competencies (Ulrich, 1997), and business competencies (Brewster, Farndale, & Ommeren, 2000).

One of the reasons for the lack of a generally accepted model of competencies was addressed by Schippmann et al. (2000). The authors posited that there is no agreed upon definition of competency models— this creates issues for creating standard measures. Some definitions of competencies may be more focused on the knowledge, skills, and abilities of a position, whereas others may be more closely related to trait-oriented characteristics, such as personality (Schippmann et al., 2000). Concern exists among researchers that competencies closely related to personality traits may present problems for practitioners (Latham & Wexley, 1994). This unease subsists because competencies are seen as being somewhat changeable, whereas personality characteristics may be seen as predispositions that are relatively stable over time (Schippmann et al., 2000). An illustration of the problem with competencies being too closely related to personality characteristics is addressed by Peter Drucker, stating:

An employer has no business with a man's personality. Employment is a specific contract calling for specific performance and nothing else. Any attempt of an employer to go beyond this is usurpation. It is immoral as well as illegal intrusion of privacy. It is abuse of power. An employee owes no "loyalty," he owes no "love," and no "attitudes," he owes performance and nothing else... Management

and manager development should concern themselves with changes in behavior likely to make them more effective (Drucker, 1973; pp. 424-425).

A dichotomous categorization of competencies in trait-based and skill-based competencies was presented by Spencer and Spencer (1993). The authors stated that skill-based competencies are comprised of knowledge (i.e., information or expertise in an area), skills (i.e., behavioral demonstration of expertise), and motives (i.e., recurrent thoughts that drive behavior); whereas, trait-based competencies are made up of self-concepts (i.e., attitudes, values, and self-image) and traits (i.e., the general dispositions of a person; Spencer & Spencer, 1993; Vazirani, 2011). Banasova, Caganova, and Cambal (2011) defined trait-based competencies as those abilities focused on the individual emphasizing how something is achieved and what individual characteristics that may be necessary for an individual to accomplish a task of a required level. The researchers also defined skill-based competencies as abilities based on work tasks and behaviors with an emphasis on what should be achieved and the behaviors that must be carried out for task completion. Trait-based competencies are given the distinction as being *competencies*, while skill or behavior-based competencies are distinguished as *competence* (Banasova et al., 2011; Vazirani, 2011). Previous research has seen no issues with the two types of competencies (Spencer & Spencer, 1993), as long as each competency is related to performance in the workplace (Banasova et al., 2011; Vazirani, 2011).

Some 360-degree feedback processes are based on the development of competencies (Caputo & Roch, 2009). However, the literature on competency modeling can be challenging to interpret. This is due to the lack of uniformity in competency model

definitions, different purposes and goals of the feedback process, and the use of 360-degree feedback with multiple job levels (Caputo & Roch, 2009; English, Rose, & McLellan, 2009). Competency models do have many advantages. They can clarify work expectations, create shared understanding of expectations among individuals, and serve as measure of human capital that an organization possesses (i.e., talent management) or wants to strategically possess (i.e., workforce planning) to gain competitive advantage (Latham & Wexley, 1994). Consequently, Schippmann et al. (2000) stated that people with the right competencies, or people who have the potential to develop the right competencies, will be more likely to elicit the appropriate behaviors to produce the desired organizational or personal results.

Within 360-degree feedback mechanisms, the dimensions on which managers are rated tend to be comprised of competencies. Banasova et al. (2011) stated that one of the most frequently used applications of competency models is for staff appraisal, such as 360-degree feedback, and that this method is used in many different contexts. Using 360-degree feedback, competency models can be incorporated to develop individuals and help reduce the gap between competencies required for success, and existing capabilities (Beheshtifar & Moghadam, 2011; Rothwell, 2005). According to Abdullah et al. (2011), competency models allow for more strategic human resource development for a long term issues by mitigating changes in society, industry, economic conditions, legislation, globalization, and technological issues. Utilizing 360-degree feedback constructed from competency models may help organizations stay competitive in the twenty-first century, where business rapidly changes and the development of robust leaders is essential (Vazirani, 2011).

### **Overview of a 360-Degree Feedback System**

To understand how 360-degree feedback competencies can be used for development, a person must first understand how 360-degree feedback systems are built and implemented (Goldsmith & Underhill, 2001). In its simplest form, all 360-degree feedback systems share a number of common elements. These elements include: the reason for completing the assessment (i.e., employee development), the person being assessed (i.e., the ratee), the persons making the assessments (i.e., the raters), specific areas being measured (i.e., managerial competencies), techniques for data collection (i.e., survey instruments), methods of interpreting rater responses (i.e., analyzing the data), a means to convey the results (i.e., feedback report), and a person to provide the results (i.e., direct feedback) who will presumably change behavior (Fitzgibbons, 2003). The link between 360-degree feedback and improved performance is clear (Seifert et al., 2003), and there should be an emphasis on the importance of developmental activities following the feedback session (Goldsmith & Underhill, 2001).

Ninety percent of Fortune 500 companies have used some form of multi-rater feedback (Maylett & Riboldi, 2007); with some of these feedback programs developed internally, while others are implemented by external groups (e.g., industrial and organizational psychology consulting organizations). Within these interventions, almost all of the managers who have participated in the developmental feedback have found the feedback to be helpful (Wood et al., 2006). When conducted on multiple occasions, seventy-six percent of executives participating in 360-degree feedback assessments were rated as more effective leaders after at least six months after they received the initial feedback (Goldsmith & Underhill, 2001).

However, not all organizations benefit from managerial improvement after 360-degree feedback. Managers may fail to improve because 360-feedback interventions are not implemented properly. According to London and Smither (1995), organizations implement multi-source feedback in different frequencies: (a) 40 percent only administer it once, (b) 25 percent administer it twice, (c) 15 percent administer it annually, and (d) 20 percent administer it once, then again on irregular intervals. The inconsistency of organizations in their use of 360-degree feedback can be a hindrance to the overall effectiveness of the system.

In a broad sense, multi-source feedback is a mechanism of introducing culture change (London & Smither, 1995); however, it has potential risks. The cost of implementing a multi-source feedback program tends to be quite expensive and the system may take up a considerable amount of supervisor and subordinate time (Morgeson et al., 2005). With such a large investment of time and resources put into the program, it is critical that organizations know whether such a program leads to improvement and whether the improvement can be sustained over a long period of time (Dai, De Meuse, & Peterson, 2010). Aguinis (2008) mentioned that for management interventions to be successful, they must include the following attributes: congruence with the organization's mission and vision, thoroughness, practicality, reliability and validity, meaningfulness to participants, and be considered fair and equitable. Moreover, being aware of the process needed for an intervention to be successful is important, but there are strategic considerations for implementing an intervention.

360-degree feedback systems are typically structured the same across organizations and industry (Fitzgibbons, 2003), even though the purposes of the

interventions may differ. Morgeson et al. (2005) indicated that there are over 20 dimensions of job performance and development that should be considered when implementing 360-degree feedback. When designing a 360-degree feedback intervention, ratings may be made on performance dimensions strategic to organizational success and relevant to the job. The typical starting point for these interventions comes from job analysis to determine the appropriate content for the assessment (London & Beatty, 1993). Also, in many instances performance dimensions are outlined through the job analysis based on competency models (Brannick et al., 2007). As a part of the 360-degree feedback process, managers can be rated on dimensions of their behavior and performance on which they may not have been previously rated (London & Beatty, 1993). When defining the content of the behavioral dimensions, it is important to involve a group of knowledgeable employees to help identify and generate behavioral statements (Brannick et al., 2007). Items can be general or specific and should reflect prototypical managerial behaviors or competencies (London & Beatty, 1993). The more an item represents a behavior or competency, the more salient the item will be to the rater (Cronshaw & Lord, 1987). When including a job analysis and aligning the intervention with organizational goals, developmental 360-degree feedback programs have a higher probability of increasing performance.

One of the steps in creating a multi-source feedback tool is identifying the raters. Greguras and Robie (1998) suggested the optimal number of raters to achieve acceptable levels of reliability (i.e., .70) would include at least four supervisors, eight peers, and nine direct reports. Yet, in real world settings, that number of raters may rarely be practical or feasible. According to Aguinis (2008), 360-degree feedback systems require ratings from

supervisors, peers, subordinates, self, and customers. However, the author stated that the supervisor has the ultimate responsibility for managing the rating process and ensuring the manager's evaluation of performance coincides with organizational strategy. When providing training and instructions to raters, the training should be unambiguous and explain the purpose of the rating process, how the data will be aggregated, and how results will be fed back (London & Beatty, 1993).

Feedback ratings are often accompanied by managers' self-ratings on the items on which they were also rated by their supervisor, subordinates, peers, and customers. Self-ratings help focus the manager's attention on the results and build motivation in establishing the direction of self-development efforts (Meyer, 1991). After all the ratings are collected, analyzed, and synthesized in a logical manner, a developmental report can be compiled for the feedback recipients' use (Fitzgibbons, 2003). The developmental report can include a narrative statement summarizing the results, item-by-items listings for each of the rater groups, and data averages across predetermined item factors (London & Beatty, 1993). With the final report complete, the first step in the developmental feedback process is underway and the feedback session can be arranged.

During the 360-degree feedback session, the developmental report is discussed with the manager, including the competencies identified as strengths and areas identified for improvement (Fitzgibbons, 2003). In the feedback process, sometimes managers receive the feedback report for self-interpretation, and other times there is a feedback facilitator meets with the manager either individually or collectively to review the feedback report. The facilitator is usually the supervisor of the 360-degree feedback participant or a member of the Human Resources staff (Seifert et al., 2003). Feedback

recipients typically select to develop their bottom-ranked competencies in managerial performance feedback programs (Dai et al., 2010). It is sometimes considered less threatening to have a feedback facilitator who is a consultant or human resource professional than to have a superior responsible for evaluating the manager's performance (Antonioni, 1996). However, the person giving the feedback is not the only important element in the feedback intervention because sometimes feedback is given on multiple occasions.

If 360-degree feedback interventions are made on multiple occasions, improvement can be measured or performance deficits can be monitored and coached (Dai et al., 2010). However, the areas where feedback is given and coached are not only related to task improvement, but also to competencies. This can be a problem because competencies can be broken into trait-based and skill-based approaches (Vazirani, 2011). Competencies that are more closely related to traits will be less likely to be developed over multiple iterations due to the stable nature of traits (Spencer & Spencer, 1993).

According to Atwater and Brett (2006), managers who receive feedback in a numeric format react more favorably than those who receive text feedback, regardless of the source of feedback. These findings suggest that feedback providing scores and comparative information is received more positively than text feedback providing only self-relevant data. Numeric feedback can be discussed between the supervisor and manager, and gaps in the ratings can be identified as growth areas, such as gaps between rater types or selected competencies (London & Beatty, 1993). The structure and execution of the feedback session is essential for the success and development of

managers using 360-degree feedback. A strategically aligned 360-degree feedback system can sustain a competitive advantage for participating organizations (Vazirani, 2011).

### **Current 360-Degree Feedback Research**

Research on 360-degree feedback continues to be an area of interest in the industrial and organizational psychology purview. According to Morgeson et al. (2005), more than 100 articles relating 360-degree feedback have been published since 1990. During that time, it has been reported that multi-source feedback is being used by the majority of Fortune 1000 companies in the United States, Australia, and around the world (Carruthers, 2003). Multi-source feedback systems have proliferated throughout business and are being used for diverse purposes and interventions, such as executive coaching, performance evaluation, talent management, succession planning (Morgeson et al., 2005). According to Nowack and Mashihi (2012) 65 percent of small businesses use some form of multi-source feedback for their workers. On the basis of 13 longitudinal studies, evidence supports the improvement of managerial performance over time (Smither, London, Flautt, Vargas, & Kucine, 2002). This finding supported the wide use of 360-degree feedback in business; however, the effectiveness of 360-degree feedback still largely depends on the format and execution (London & Beatty, 1993).

In a five-year longitudinal study of participants in a multi-source feedback program, a high level of internal consistency ( $\alpha > .90$ ) was found for ratings over time, providing evidence that 360-degree feedback ratings have a high level of reliability and are stable (Violato, Lockyer, & Fidler, 2008). Contrary to this, Kluger and DeNisi (1996) proposed several reasons why feedback may not be effective or stable over time,

including characteristics of the feedback, the task, and the recipients. Feedback recipients are not always the best judges of their own strengths and weaknesses (McPherson, 2007), and one way of finding out how others view their performance is through the use of 360-degree feedback evaluations.

Contrary to previous findings of self-rating inflations, Hassan and Rorhbaugh (2006) found that situations do arise in which managers underestimate, rather than overestimate, their own performance. This could be an artifact of managers' roles in the organizations changing over time, or the personalities of people selected for management roles may have changed (Smither et al., 2003). However, similar to previous research, self-ratings are still the most inaccurate forms of rating included in 360-degree feedback, whether inflated or deflated (Hassan & Rohrbaugh, 2006). In their research on self-ratings, Van Velsor et al. (1993) found that over-raters (e.g., those with self-ratings above other-ratings) received the lowest subordinate ratings on managerial roles compared to under-raters (e.g., those with self-ratings below other-ratings) who received higher subordinate ratings.

These results are consistent with the current socio-psychological research, indicating that highly competent people will underestimate their own performance (Burson, Larrick, & Klayman, 2006; Krugar & Dunning, 2002). Mersman and Donaldson (2000) suggested that, because under-raters may exhibit a lack of self-confidence, they may have more difficulty in making important decisions regarding their future career goals and objectives. This is important because although under-rating managers may be rated higher by others, they may not actually be better leaders or managers (Mersman & Donaldson, 2000). The effect of self-rating incongruence has been debated in 360-degree

feedback research for decades, but the present research proposed that the validity of self-ratings could be tested. However, another current issue related to 360-degree feedback research is the applicability of its interventions across cultures around the world because of the globalization of work (Pollitt, 2005).

### **Culture and 360-Degree Feedback**

As more day-to-day business is conducted on an international basis, culture will play an increasingly important role in business, specifically when referencing the current status of research on 360-degree feedback. Hofstede's (1980) five cultural dimensions (e.g., Masculinity/Femininity, Power Distance, Long Term Orientation, Individual/Collectivism, and Uncertainty Avoidance) have been considered as having possible moderating effects on 360-degree evaluations (Hofstede, 2001). Hofstede (2001) found that high Power Distance and low Power Distance cultures may play a moderating role on 360-degree feedback effectiveness, defined by countries Power Distance Index. Eckert, Ekeland, Gentry, and Dawson (2010) also suggested that systematic differences in self- and other-ratings may due to cultural variables, such as Power Distance. Hofstede (1980) defined Power Distance as the extent to which a community accepts and endorses authority, power differences, and status privileges. Although differences in cultures among the dimension of Power Distance are the only found in current literature, issues of cross-cultural impact of 360-degree feedback systems have not been fully addressed (Mittal & Saran, 2010).

Hofstede (2001) found that 360-degree feedback systems are only effective across differing cultural dimensions if the feedback intervention is accepted by employees of that culture and if the system is confidential while not violating cultural norms.

According to Eckert et al. (2010), self-other differences in ratings are present in cultures where observers have a more distant perspective on leaders. In these cultures, the discrepancy between self and observer ratings was found to be larger (Carl, Gupta, & Javidan, 2004). On a global scale, business interventions such as 360-degree feedback may be changing the global culture and opinions toward authority (Eckert et al., 2010). However, cultural dimensions are not the only current issues confronting 360-degree feedback researchers. The degree of performance improvement using 360-degree feedback interventions may also play a role in 360-degree feedback research.

### **Positive Organizational Outcomes Using 360-Degree Feedback**

Ratee reactions to feedback have been linked to the effectiveness of 360-degree feedback and are considered immediate predecessors to performance improvement (Bailey & Austin, 2006; Chappelow, 2004; Leslie, 2002; Maurer et al., 2002; Richardson, 2010; Wood et al., 2006). When 360-degree feedback is appropriately planned and executed, it can enhance team-working, productivity, communication, and trust (Wood et al., 2006). However, researchers continue to debate whether 360-degree feedback is the panacea of performance improvement that its supporters claim it to be.

Because of the inclusive nature of 360-degree feedback, the person being rated, the raters, and the organization all stand to benefit from an increase in employee voice and performance improvement from managers (Wood et al., 2006). Through the engagement that comes from 360-degree feedback participation, Maylett and Riboldi (2007) found that employees provide better customer service, record lower rates of attrition and absenteeism, demonstrate improved quality, and exhibit increased productivity, all of which are related to overall performance. Multi-source feedback

provides an opportunity for a manager to demonstrate support and personal investment in the participatory structure of the feedback, also allowing managers to become better role models by using the process (Chappelow, 2004; Hernez-Broome & Hughes, 2004; Richardson, 2010). With the help of a supportive facilitator, recipients are more likely to set improvement goals and develop plans for improving their performance (Bracken, 1994; DeNisi & Kluger, 2000; Seifert et al., 2003).

Bailey and Austin (2006) posited that participation in 360-degree feedback can lead to positive and negative outcomes for focal individuals. The authors stated that variability can be attributed to three factors: (a) whether the feedback is received favorably, (b) whether the participant's self-efficacy is supported or challenged, and (c) the perceived importance of changing the behaviors on which feedback is presented. Extremely negative feedback can lead recipients to abandon their goals to perform more effectively (Kluger & DeNisi, 1996), and high disagreement between self- and other-ratings can lead to lower performance (Ostroff et al., 2004). There is an increasing need for strategic planning for 360-degree feedback assessments, along with the appropriate training and support staff for interventions to be effective (Bailey & Austin, 2006).

Ten years after London and Smither's (1995) seminal work on the performance benefits from 360-degree feedback, the researchers stated, in opposition to their earlier recommendations, that practitioners should not expect large, widespread performance improvements after employees receive multi-source feedback (Smither, London, & Reilly, 2005). If feedback systems are executed appropriately, 360-degree feedback can

assist in building a collaborative and participative organizational culture (Maylett & Riboldi, 2007; Richardson, 2010; Wood et al., 2006).

One possible effect 360-degree feedback could have on a manager is increasing the manager's self-awareness and self-perception. Multi-source feedback may be used as information-gathering tool from multiple sources serving as a vehicle for self-awareness, assessment, and development (Bliszczyk & Dimasi, 2003; Bracken & Timmrick, 2001; Leslie, 2002; McCarthy & Garavan, 2001; Richardson, 2010; Wood et al., 2006).

Self-awareness can be defined as an aspect of the personality related to the awareness of a person's individuality and his or her relationship to others in interpersonal relationships (Wood et al., 2006). The extent of the congruence of self- with other- ratings has been used as a measure of self-awareness, and this personality variable has been found to be significantly correlated to effective performance outcomes (Fletcher & Baldry, 2000).

Gaining self-insight and a broadened perspective seems to be a general theme in the effectiveness of developmental feedback interventions for managers. Research on the relationship between self-awareness and 360-degree feedback (Bracken & Timmrick, 2001; Wood et al., 2006), has shown that leaders with higher levels of self-awareness (i.e., self-other agreement) tended to be better leaders and more responsive to 360-degree feedback interventions. Hagan et al. (2006) found additional evidence for the importance of self-awareness, in which significant correlations between 360-degree assessments of core competencies were related to assessment center performance ratings. Nevertheless, self-awareness may not be the only personality construct with a relationship to performance on 360-degree feedback interventions. Smither, London, Flautt, Vargas, and Kucine (2003) investigated the link between 360-degree feedback and participation in

executive coaching. The authors found the subsequent impact on 360-degree feedback ratings had high variability in feedback outcomes, which could be attributed to individual differences or situational variables. This finding led to the hypothesis that the effectiveness of 360-degree feedback may be mediated by individual differences or personality characteristics. The relationship between personality and 360-degree feedback remains a current topic in the research.

### **Personality and 360-Degree Feedback**

Although some recent topics of research on 360-degree feedback include elements of culture and the impact of organizational outcomes, one area needing more research is the relationship of personality to 360-degree feedback competencies (Smither et al., 2003). While some researchers have criticized the use of personality variables in explaining behavior at work in the past (Mitchell, 1979), personality variables have been found to predict job performance in organizations (Weiss & Adler, 1984). In previous research, assessments of an individual's strengths and weaknesses have been termed self-awareness by a number of researchers (London & Smither, 1995; Wohlers & London, 1989), and the concept of self-awareness was the first personality construct researched in relation to 360-degree feedback (Church, 1997). However, more recently, new personality measures have become more adept at analyzing the relationship of personality to a person's job. One of the most common personality correlates of performance is called the Five Factor Model (FFM; Hogan, 2004).

#### **Five Factor Model**

Research on personality was a growing area in the early twenty-first century in terms of the job-relatedness literature, specifically research on the *Big Five* or Five Factor

Model (Mitchell & Daniels, 2003). No single researcher has been credited with the development of the Five Factor Model, and numerous studies have come to the same five factor structure independently (John, Naumann, & Soto, 2008). The five personality variables associated with the Big Five include Neuroticism, extraversion, conscientiousness, Agreeableness, and Openness to Experiences; although other terms for each of the variables have been used interchangeably (Costa & McCrae, 1992). Currently, the most commonly used measure of the Five Factor Model of personality is the NEO Personality Inventory (McCrae, Costa, & Martin, 2005). The present research discusses the implications of the Five Factor Model (FFM) on personality and job performance.

In their meta-analysis of the relationship between personality and managerial success, Judge, Bono, Ilies, and Gebhardt (2002) found significant direct effects for extraversion (.31), emotional stability (.24), agreeableness (.08), conscientiousness (.28), and openness to experiences (.24). From this meta-analysis, it appeared that all the factors in the FFM may be related to managerial success. When being marketed to potential users, 360-degree feedback appraisal competencies tend to claim a developmental link to managerial success (London & Smither, 1995). From claims such as these, it is not difficult to infer that possible linkages exist between the FFM, ratings on 360-degree feedback mechanisms, and managerial performance. Even though previous research has shown that all Big Five personality traits are significantly related to managerial success, certain personality constructs of the FFM may be more important to job performance than others.

Leaders or managers that are high in extraversion, conscientiousness, and emotional stability have been shown repeatedly throughout the literature to have higher ratings of work performance (Barrick & Mount, 1991; Gray, 1994; Smither, London, & Richmond, 2006). In regards to multi-source feedback, the personality trait of extraversion has been shown to be positively related to requesting additional feedback in managers, and conscientiousness has been shown to be positively related to subsequent participation in developmental activities (Smither et al., 2006). According to Barrick and Mount (1991), extraverted leaders tend to have higher performance ratings than those whom are not extraverted. Conscientiousness has been found to be related to setting and attaining goals after receiving peer feedback (Dominick, Reilly, & Byrne, 2004), and conscientious leaders are also more likely to use the feedback they received when participating in 360-degree feedback interventions (Smither et al., 2006). Dominick et al. (2004) found that conscientiousness along with openness to experience was positively related to the performance of managers after receiving peer feedback. In terms of emotional stability, Atwater and Brett (2006) found that leaders with low emotional stability reported more negative emotions (e.g., angry, frustrated, unhappy, discouraged, and disappointed) after receiving feedback, even though they do not receive less favorable feedback than other leaders.

### **Other Personality Perspectives**

It is important to note that other relationships between personality and 360-degree feedback competencies have been researched outside of the FFM. Smither et al. (2005) found that some feedback recipients are more likely to improve than others because of their individual differences. Personality traits, such as levels of self-efficacy, belief in

human ability, regulatory focus, and emotional intelligence have been found to moderate the effectiveness of 360-degree feedback (Atwater & Brett, 2006; Bailey & Austin, 2006; Funderburg & Levy, 1997; Leslie, 2002; Maurer et al., 2002; Richardson, 2010). In particular, self-efficacy has been shown to be important to managerial improvement following 360-degree feedback (Atwater & Brett, 2006; Bailey & Austin, 2006; Richardson, 2010). The theoretical and empirical rationale for the effect of self-efficacy on post-feedback behavior derives from the literature on Social Learning Theory (Bandura, 1977), Goal Setting Theory (Locke, 1996), and Feedback Impact Theory (London & Smither, 1995).

Using feedback impact theory, London and Smither (1995) identified several individual difference variables that can affect interpretation of feedback, including self-image, feedback-seeking behaviors, and self-monitoring. The authors explained that positive reactions were greater for feedback recipients with high self-efficacy, and lower for feedback recipients who focused more on managing others' impressions of them rather than improving behavior. Overall, self-awareness is one of the primary outcomes of 360-degree feedback (Atwater & Brett, 2006; Bailey & Austin, 2006; Goldsmith & Underhill, 2001; Richardson, 2010).

Other individual difference variables have been found to effect 360-degree feedback ratings. Ostroff, Atwater, and Feinburg (2004) studied of over 4,000 managers across 650 organizations and found that individual differences exist for many different reasons. The researchers described differences between self- and others-ratings related to gender (e.g., men were rated less favorably by others than were women), race (i.e., non-whites rated themselves higher than whites), age (i.e., older managers rated

themselves higher but were rated lower by others), experience (i.e., experienced managers rated themselves higher than less experienced managers), and education (i.e., managers with less education were rated lower by others but did not differ in self-ratings from managers with more education). Finding rating incongruence, such as this, indicates the need for more research into individual differences in ratings and the comparison to multi-source feedback. The current study intends to look into the relationships of personality and individual difference variables and 360-degree feedback competencies to fill this gap in the literature.

### **Measures of Personality**

The Guilford-Zimmerman Temperament Survey (GZTS) is one of the most widely used personality inventories available, and has acceptable levels reliability and validity (Guilford, Zimmerman, and Guilford, 1976). The Dynamic Factors Opinion Survey (DFOS) was also developed by Guilford (Guilford & Martin, 1944) and its focus is on the assessment of motivational and human needs (e.g., need for attention). As with the GZTS, the DFOS has been found to have reliable and valid interpretation (Guilford, Christensen, & Bond, 1956). Using factor analysis, Guilford and Martin (1944) derived ten factors from the DFOS, yet within the combined GZTS and DFOS measure, more of the GZTS scales were used than the DFOS scales. All scales on the GZTS and DFOS have had internal reliability coefficients above .70 in previous research, which according Nunnally and Bernstein (1994) is considered an acceptable level of internal consistency. In their attempt to merge both instruments, a private consulting firm, using a system of expert judgment, combined the 300 items GZTS and the 300 item DFOS to create a 350

item personality assessment (GZTS/DFOS personality inventory; Technical Manual, 2009).

Glasgow (1999), as a part of her dissertation research, examined the correlations between the GZTS/DFOS personality inventory and the NEO PI-R (based on the Five Factor Model). In her sample of 88 professionals, many significant relationships were reported between the two measures, and she concluded some overlap does exist between the dimensions measured by the Big Five and those measured by the combined GZTS/DFOS personality inventory (Glasgow, 1999). The relationships between the two measures are found in Table 1.

Table 1

Relationship Between the GZTS/DFOS and the NEO-PI-R

NEO-PI-R	GZTS/DFOS
Neuroticism	Optimism (-); Emotional Evenness (-)
Extraversion	Sociability; Assertiveness; Work Pace
Agreeableness	Need to be Liked; Positive about People
Conscientiousness	Self-Reliance; Work Pace; Serious-Minded
Openness to Experience	Liking for Thinking, Detail Interest

*Note.* From Glasgow, L. P. (1999). A comparison of broad and narrow personality traits in the prediction of job performance. Unpublished Doctoral Dissertation.

Within Glasgow's (1999) research, regression analyses were also conducted to predict the Big Five. She found that the NEO-PI-R did not add any incremental predictive validity beyond the combined GZTS and DFOS, but the GZTS/DFOS personality inventory did add predictive incremental validity above the NEO-PI-R variables. Her finding is important because it supports the theory that the combined GZTS/DFOS personality inventory predicts the Big Five factors just as well as the NEO-PI-R, and

possibly even better. In terms of the current study, elements of the GZTS/ DFOS personality inventory are being used as a surrogate for NEO-PI-R scales. The scales of the GZTS/DFOS personality inventory serve as an operationalization of the FFM used to illustrate some of the relationships between personality (i.e., the Big Five) and 360-degree feedback competencies.

### **Models of Personality and 360-Degree Feedback**

Currently, few models of the personality and 360-degree feedback relationship have been posited by researchers. The majority of prediction models assumed linear relationships between each personality trait and performance criteria (Lin, 2012). Models for predicting 360-degree competency ratings may have failed to capture all of the unique variance in ratings using only simple regression. Latham and Wexley (1994) stated that some of this unique variance may be missing because of the aggregation techniques used in the process of analyzing 360-degree competency data.

One research study on the linear relationship between personality and 360-degree competency-based ratings was conducted by Lin (2012). The sample consisted of 804 directors or senior managers. The Occupational Personality Questionnaire (SHL, 1999) was used to measure personality and the Inventory of Management Competencies (SHL, 1993) was used to measure 360-degree feedback competencies. The analyses were only conducted on two levels: (a) single trait to single rater and (b) multiple personality traits were regressed towards aggregated 360-degree feedback competency dependent variables. Lin's (2012) findings showed little evidence for uni-dimensional linear relationships from aggregated personality scores to averaged 360-degree feedback scores. To clarify the aggregation techniques used for 360-degree feedback, Latham and Wexley

(1994) stated that currently, when using aggregation or averaging techniques, several steps are involved. First, ratings between raters (i.e., subordinates, peers, supervisor, self, and others) are aggregated or averaged on each item. Second, item-level ratings are aggregated or averaged to give a combined rating per 360-degree feedback competency (Latham and Wexley, 1994). After aggregating or averaging ratings, then techniques, such as simple regression, can be used for 360-degree feedback competencies to serve as criteria in linear models. This research hypothesizes that, due to the large amount of variance left unobserved by aggregating or averaging procedures, other methods of analyzing 360-degree feedback competencies may be more effective than the current methods.

Seifert et al. (2003) created a model of 360-degree feedback in which feedback orientation and personality moderate the relationship between the initial reaction to feedback, goal setting, taking action, and performance improvement. The authors suggested that future models include a more comprehensive framework for performance improvement and multi-source feedback, in which individual differences are included, such as personality, goal setting, and feedback orientation (Seifert et al., 2003). Despite the expanded coverage of the model, the revised model did not imply any causal paths between personality and 360-degree feedback competencies. The current study plans to expand on Seifert's et al. (2003) notion that individual difference variables may affect 360-degree feedback competencies.

Not all researchers believe that personality is related to outcomes in 360-degree feedback. Richardson (2010) and Brusman (2008) proposed models of 360-degree feedback in which personality traits and styles are not addressed during the process.

However, when classifying dimensions of manager performance, Yukl and Van Fleet (1990) stated that performance should consist of three categories: administrative (e.g., planning, organizing), human relations (e.g., working with and through people to accomplish objectives), and technical competence (e.g., knowledge of relevant or new techniques). The present research proposed the rationale for not including personality as a component of 360-degree feedback competencies. This is derived from the view that some competencies are based on relatively stable traits (i.e., trait-based competencies) and may not be changeable or learnable attributes that can be improved by developmental activities. The observation that personality traits are relatively unchangeable and possibly related to 360-degree feedback competencies has raised concern for researchers advocating the use of 360-degree feedback for manager development (Richardson, 2010). The concern is that personality traits amenable to change (Richardson, 2010). The present research proposed that personality traits of the Big Five (Costa & McCrae, 1997) are related to each of these dimensions of 360-degree feedback competencies (i.e., trait-based competencies).

### **Structural Equation Modeling**

There may be an inherent suitability in using latent models of performance in the examination of 360-degree ratings because of the similarities between construct-item relationships in structural equation modeling (SEM; Bagozzi, 1994) and competency-item relationships in 360-degree feedback (Vazirani, 2011). However, very few studies have investigated the construct validity of different stakeholder groups in 360-degree feedback using SEM (Silvester & Wyatt, 2012). Latham and Wexley (1994) described construct validity as a condition for establishing where job-relatedness of an

appraisal system is true and meaningful, which may be used if other means of validity are not technically possible. A high intercorrelation of ratings, as well as convergent validity, is one indication of construct validity (Latham & Wexley, 1994). Previous research has acknowledged that multi-source feedback only has a moderate level of inter-rater agreement (Harris & Schaubroeck, 1988). Depending on the method used to aggregate or average ratings from different sources, 360-degree feedback ratings may be combined in ways to create intercorrelations high enough to imply construct validity (Lawler, 1967).

Two ways of analyzing the construct validity of a measure are using confirmatory factor analysis (CFA) and SEM (Steenkamp & Baumgartner, 2000). An important consideration when using both CFA and SEM is sample size, and Bentler (1985) suggested a sample size to parameter ratio of five or more may be sufficient to achieve reliable estimates. Taking a parsimonious approach when developing scales used in SEM is one strategy for obtaining a stable or reliable solution, and this can be accomplished by developing scales that attain a high level of internal consistency with few items (Van Velsor, 1998).

### **The Theoretical Basis of SEM**

Structural equation models make it possible to identify measurement error and mathematically correct for attenuation due to measurement error (Steenkamp & Baumgartner, 2000). The philosophy of theoretical constructs is operationally defined by Bagozzi (1994) as the one-to-one correspondence between a theoretical construct and the constructs' measurement. The author noted that, in practice, most modelers implicitly assume the observed variables are perfect measures of the underlying constructs. This essentially precludes any meaningful distinction between the construct and the

constructs' particular operationalization. It can be risky for a researcher to assume that the operationalization of a construct is actually equal to the implicit construct being measured (Bagozzi, 1994). That is why, for construct validity to be inferred, research findings should be assembled over time to create a body of research supporting the validity of the construct (Uniform Guidelines, 1978).

Steenkamp and Baumgartner (2000) point out that when researching the relationship between two constructs, sometimes the relationship is not found to be supported empirically. Whether there is truly no relationship is unclear. The relationship may be masked by measurement, or the variables may lack validity because they fail to measure what they claim to be measure (Steenkamp & Baumgartner, 2000). One reason SEM can be useful is that it makes a clear distinction between observed, theoretical constructs and fallible, empirical measures. Steenkamp and Baumgartner (2000) stated that this is based on the partial interpretation philosophy, which advocates a doctrine of multiple operationalizations of the underlying construct by individually imperfect but collectively reliable and valid measures.

Bagozzi (1994), in his seminal work on SEM, stated that before the measurement model can be compared to the latent model using SEM, a CFA of the model parameters must be conducted. According to Steenkamp and Baumgartner (2000), if several indicators of a construct are available, it is a relatively straightforward task to assess the reliability and validity of sets of indicators before they are used in a structural model. The usual procedure for specifying a CFA model is to investigate how well the multiple indicators capture the constructs of interest. Random error may be isolated using CFA, which is why it is recommended that it be used before proceeding to SEM (Hair, Black,

Babin, Anderson, and Tatham, 2006). Systematic error is more difficult to identify because it can be due to semantic issues or structural issues in the data (Steenkamp & Baumgartner, 2000).

The multitrait-multimethod approach may be used to partial out the differences between random error and systematic error (Mount et al., 1998). When using CFA for 360-degree feedback competencies, the model assumes that each variable contains method variance, trait variance, and unique variance (Conway, 1996), and this allows for determination of the degree to which raters and traits account for co-variation among measures (Mount et al. 1998). Chi-square ( $\chi^2$ ) is the statistic used to compare measurement and latent models, and the larger a significant chi-square is, the larger difference which exists between measurement and latent models (Hair et al., 2006). While using CFA, researchers attempt to create models that have small chi-squares, indicating relative similarity between measurement and latent models (Steenkamp & Baumgartner, 2000). When comparing models, Bentler and Bonnett (1980) noted that the size of the chi-square is a direct function of sample size of the data. The researchers noted that small sample sizes tended to yield non-significant chi-square statistics, even when models did not fit the data well. Due to concerns over significance testing when using the chi-square statistic and its limitations (Bollen, 1989; Schmidt, 1996), current researchers are using other methods to analyze the data in CFA models, such as standardized fit indices like the root-mean-square error or approximation (RMSEA) and the comparative fit index (CFI; Bagozzi & Yi, 2012).

Once a CFA has been conducted and obtains the appropriate fit and chi-square, the model can be transformed into an SEM path or linear model, based on theoretically

causal relationships (Bagozzi, 1994). Structural equation models are usually employed in studies investigating the structured linear relations between constructs based on cross-sectional data (Steenkamp & Baumgartner, 2000).

According to Steenkamp and Baumgartner (2000), the SEM approach to research is guided by three principles. First, many, if not most, scientific constructs have facets that cannot be directly observed, and these are sometimes referred to as constructs. Constructs can only be measured through observable measures or indicators that vary in their degree of observational meaningfulness and validity. No single indicator can capture the true theoretical representation of the underlying construct and hence, multiple indicators are necessary. Second, observed measures of theoretical constructs are invariably contaminated with measurement error, and the correspondence between constructs and their measures needs to be explicitly stated by the model. Third, models are always simplified representations of reality and before any conclusions are derived from a model, the degree to which the model is in agreement with the data must be ascertained.

By convention, Greek letters are used to depict parameters estimated, circles to represent latent constructs, and boxes to indicate item measures (Farh & Dobbins, 1989). Each construct is measured by multiple indicators so that measurement error can be taken into account (Hair et al., 2006). Usually, the model of interest consists of several equations describing the interrelationships among several endogenous and exogenous variables. The SEM methodology tests the equality of structural relationships, and as long as the model remains identified (i.e., over three items per construct and large number of degrees of freedom) and assumptions of independence of errors are met, errors in

equations are allowed to correlate (Bagozzi & Yi, 2012). SEM models that are based on valid constructs, lacking spurious relationships, and tested repeatedly over time can make claims of construct validity (Hair et al., 2006).

### **Construct Validity**

Brannick et al. (2007) described the term construct (which can be used in both confirmatory and structural models) as being operationally defined as the underlying psychological factor that an assessment or test is claiming to measure. The *Uniform Guidelines of Employee Selection Procedures* (1978) states that, when determining construct validity, one should portray evidence that the trait or construct being measured is important to success on the job. The construct may not be directly measureable (i.e., intelligence), but methods of aggregating and compiling behaviors and characteristics of the construct can be made over large sample sizes to provide evidence that the construct may, in fact, exist and have predictive validity for the job (Brannick et al., 2007). Construct validity of performance measures is used to infer the degree to which the persons being evaluated possess some quality or construct (i.e., employee worth to the organization) presumed to be reflected in the performance measure (Blum & Naylor, 1968). The procedure for determining construct validity of a performance measures involves gathering several different performance measures that logically appear to measure the same construct (e.g., intelligence) and then observing the relationship among these appraisal measures (Brannick et al., 2007). The current study plans to utilize elements of construct validity to test the relationships of personality constructs and global management competencies measured through 360-degree feedback.

Another standpoint from which to view construct validity, posited by Hair et al. (2006), is the accumulation of other types of validity evidence and the extent to which a set of measured variables actually represent the theoretical latent construct they are designed to measure. The researchers argued that construct validity is made up of four components: (a) convergent validity, (b) discriminant validity, (c) nomological validity, and (d) face validity. Convergent validity is the extent to which indicators of a specific construct converge or share a high proportion of variance in common, and this can be examined by assessing construct loadings, variances extracted, and construct reliability. Discriminant validity is the extent to which a construct is truly distinct from other constructs. Nomological validity is tested by examining whether or not the correlations between constructs in the measurement theory make sense. Face validity is the extent to which the content of the items is consistent with the construct definition, based solely on the researcher's judgment. Due to some of the limitations within the data collection, the present research looked at the convergent and discriminant validity of the constructs being measured. Nomological and face validity were not accessible to the researcher because of prior confidentiality agreements.

### **Composite and Multiple Criterion**

One of the issues with previous research on 360-degree feedback competencies is the debate over whether to analyze competencies using composite or multiple criteria. Latham and Wexley (1994) reviewed the controversy over composite versus multiple criterion measures, noting that advocates of composite criteria believe that measures should be aggregated in some manner to create a single dependent variable (Blum &

Nailor, 1968), whereas advocates of multiple criterion measures believe that criteria should be treated as multiple dependent variables (Schmidt & Kaplan, 1971). Latham and Wexley (1994) hypothesized that the use of multiple sources of criteria increases the probability of obtaining a comprehensive picture of an employee's total contribution to the organization.

An important early perspective for measuring employee performance using multiple criteria and multiple rater perspectives was the multitrait-multirater (MTMR) approach (Lawler, 1967). In his seminal article about MTMR, Lawler proposed MTMR as an alternative to the variety of objective measures that were being touted as replacements to multiple criteria. Lawler argued that more information can be obtained about the meaning of ratings using the MTMR approach than could be obtained if a single rater or single trait was being measured. Mount et al. (1998) analyzed the method effects of raters and traits in 360-degree feedback utilizing MTMR. The findings indicated that method variance is more strongly associated with individual raters, rather than the rater-level. Individual raters may be measuring different areas of performance rather than every rater measuring the same performance from different perspectives.

Another approach for analyzing the self-other and other-other comparison of agreement between rater types is the within and between analyses (WABA) technique (Yammarino & Atwater, 1997). This technique proposes that rater agreement/disagreement can take three different forms: (a) patterned agreement (i.e., self and other scores are similar), (b) patterned disagreement (i.e., self and other scores are opposite), and (c) lacking agreement (i.e., self and other scores are not related). By

underscoring the different kinds of agreement, a researcher can have a better understanding of the theoretical reasons for rater disagreement.

Schmidt and Hunter (1996) argued that inter-rater correlations can be interpreted as reliability coefficients based on a model that treats raters as passive instruments. The researchers proposed that each rater can be considered analogous to a different item on a rating instrument, and if these raters are viewed as alternate forms of a measurement instrument, the correlation between these alternative forms can constitute an estimate of reliability. Conversely, there are those who believe raters should not be treated as interchangeable forms of a rating instrument (Borman, 1974; Murphy & Cleveland, 1995). Two reasons why researchers believe using raters as interchangeable forms of a rating instrument are flawed include: (a) raters may observe different behaviors and have differing responsibilities when completing performance ratings (Borman, 1974), and (b) it implies that measurement is a primary aspect of performance ratings within an organization (Murphy & Cleveland, 1995). Building on Schmidt and Hunter's (1996) work, the present research treated individual raters as passive instruments on 360-degree feedback assessments. Using this methodology for analyzing raters makes it possible to utilize SEM to model 360-degree feedback competency's relationship to personality constructs.

However, complications and criticisms may exist for both methods of analyzing multiple criteria and composite criteria. One reason to present multi-source feedback ratings in aggregated form is because aggregated ratings have increased reliability (Scullen, 1997). Aggregating may also reduce the potentially disruptive influence of inter-rater disagreements (Gregauras & Robie, 1998). A limitation to both of these

arguments is that they assume rating aggregation is done for small sample sizes and few raters. While generally true in practice, in research there may be circumstances in which large samples sizes and adequate numbers of raters are available. Another assumption of composite criteria is the idea of aggregating ratings by source to maintain some of the variance between raters. Bozeman (1997) stated that grouping ratings for aggregation by peers, supervisors, subordinates, and others may not always make psychometric sense, but grouping can increase psychological sense-making by ratees and increase their acceptance of ratings.

The present research was concerned with the psychometric properties associated with comparing rating sources of multi-source feedback competencies as single constructs through combining ratings by rater. Also, multi-source feedback competency items could be compared as first order constructs, while superordinate competencies could be viewed as higher-order factors. This scenario has not been presented in the literature, but the concept of higher-order factors is common in the SEM literature (Hair et al., 2006). After comparing rating sources by rater, the present research hypothesized that creating first order item factors and higher-order competency factors could be a suitable method of analyzing 360-degree feedback with SEM.

### **SEM, Personality, and 360-Degree Feedback**

Generally, 360-degree feedback researchers have recommended that feedback recipients should not focus on more than two or three leadership competencies for improvement (Antonioni, 1996). Multi-source feedback uses multiple evaluation sources, varying along hierarchical organizational levels (e.g., supervisor vs. subordinates) and distance from the ratee (e.g., self vs. customer; Mittal & Saran, 2010). To date, London

and Smither's (1995) and Seifert et al.'s (2003) theoretical models of 360-degree feedback are the only published models, and there has been little research investigating the individual differences and situational variables included in models of 360-degree feedback (Bailey & Austin, 2006). In their review of the literature, Smither et al. (2005) found that virtually all of the research studies they located which investigated performance improvements did so by comparing the average or composite performance Time One ratings (before feedback) to Time Two ratings (after feedback). However, very few studies have looked at the correlates of 360-degree feedback dimensions and individual differences in 360-degree feedback responses.

There have been only a few longitudinal studies using 360-degree feedback published in the literature, and most have obtained ratings from only a single source other than the supervisor (i.e., self-ratings compared to supervisor-ratings; Reilly et al., 1996; Walker & Smither, 1999). In their longitudinal study, Dai et al. (2010) reported the extent to which different rater groups (e.g., supervisor, peers, and self-reports) agree with each other on their ratings may influence the feedback recipients' reactions to the feedback (e.g., if there is disagreement, the feedback recipient may be confused by the ratings), which may negatively affect the motivation, attitudes, and behaviors of the feedback recipient. Also, feedback recipients were observed improving more on relatively easier-to-develop competencies than more difficult competencies, especially those that are not closely related to their self-concept (Dai et al., 2010). However, this can lead to problems if the manager is not able to develop the competencies that are in most need of development. The present research hypothesized that the reason why some competencies

may be more difficult to develop than others may be due to the competency being related to a person's individual differences or personality rather than job performance.

### **Relationship Between 360-Degree Feedback and Personality**

Several researchers have expanded the relationship between personality variables and 360-degree feedback outcomes (Nowack & Mashih, 2012; Smither et al., 2005). In their summary of the evidence around 360-degree feedback, Nowack and Mashih (2012) stated that people with the personality traits conscientiousness, extroversion, high self-efficacy, internal locus of control, and low neuroticism are most motivated to use 360-degree feedback for development. Also, Smither et al. (2005) found that leaders high in emotional stability are most likely to be motivated to use feedback results for development, extroverted managers are more likely to seek more feedback six months later, and conscientious managers are more likely to engage in developmental behaviors. Moreover, managers that are extroverted and open to experience are more likely to perceive negative feedback as valuable and seek further information about their feedback.

### **Personality and Work Behavior**

In their article on personality and multi-source ratings, Warr and Hoare (2002) postulated that personality traits can predict specific work behaviors. Research into the criterion-related validity of personality scales has demonstrated that the association between predictor and behavioral criterion is stronger when those variables are aligned in terms of their content (Mount & Barrick, 1995; Robertson & Kinder, 1993; Tett, Jackson, & Rothstein, 1991). Warr (1999, 2000) reported that when examining this finding a correlation of .54 was reported in trait-behavior relationships. In previous research, the alignment of personality traits to behavioral dimensions in multi-source rating settings

has largely been ignored by researchers (Warr & Hoare, 2002). However, the present research hypothesized that stable dispositions, such as personality traits, are likely to be conceptually linked to 360-degree feedback competencies.

Brutus et al. (1999) found that personality characteristics predicted behavior ratings and this relationship is likely to be concealed if ratings are averaged into composites. Warr and Hoare (2002) cautioned that attempts to modify behaviors on the basis of multi-source ratings may unwittingly alter personality dispositions rather than job behaviors. According to Lin (2012), competency-based 360-degree rating programs are often used in leader development programs in conjunction with some form of personality inventory. Consequently, examining the relationships between personality and 360-degree feedback could be examined and be beneficial to leader development programs. According to Schmidt and Hunter (1998), the practical economic utility of a personality assessment is directly proportional to the predictive power of the concept it is measuring. When using personality inventories in conjunction with 360-degree competencies, Schmidt and Hunter (1998) believed that it is important to show the utility of the inter-relationships among personality variables and competencies.

### **Theoretical Bases of 360-Degree Feedback**

According to Hair et al. (2006), a structural model should not be built without an underlying theoretical base. In terms of 360-degree feedback, some theories, such as Self-Regulation Theory, attempt to explain the effectiveness of 360-degree feedback interventions. Through modulation of thought, affect, and behavior, Self-Regulation Theory proposed that people are able to guide their goal directed activities over time (Karoly, 1993). Expanding on Goal Setting Theory and Social Cognitive Theory, this

methodology directs people to assess their performance (i.e., self-awareness), monitor ways in which their environment facilitates or hinders goal attainment, and to identify and administer reinforcers to work toward or to punish failing to attain goals (Kanfer, 1980). Carver (2007) points out that self-regulatory efforts often run smoothly and are unimpeded by external obstacles or personal shortcomings. Vancouver and Day (2005), in reviewing the literature on self-regulation, found that self-regulatory interventions are effective in organizational settings at increasing job performance and decreasing absenteeism. Although there has been extensive research on the outcomes of performance appraisal, little research has been spent analyzing the inter-relationships of the various outcomes of 360-degree feedback programs, such as leadership effectiveness, supervisory ratings, work unit success, and customer satisfaction (London & Beatty, 1993).

Building on Drucker's (1973) philosophy that organizations should not attempt to change the traits of a man, DeNisi and Kluger (2000) stated that feedback that is directed towards the *ideal self* is not appropriate for feedback appraisals. This research study focused on dimensions of 360-degree appraisals related to the *ought self* and the *ideal self*. The *ideal self* and the *ought self* are both components of Self-Regulation Theory. As previously mentioned, Higgins (1987) emphasized that the self-regulatory model proposed that when we focus on the self we aspire to be, we are focusing on our *ideal self*. However, when we focus on the self that others expect of us, we are focusing on the *ought self*. Feedback on the *ideal self* tends to be directed toward inborn predispositions, such as traits, and can cause a person to question the core of his or her being (Higgins, 1987). This type of feedback can become problematic for 360-degree feedback because some of the competencies that make up 360-degree feedback appraisals are trait-based

(Schippmann et al., 2000). This research hypothesized that trait-based competencies are closely related to personality traits and will be more focused on the *ideal self*, while skill or behavior-based competencies are less closely related to personality traits because they are focused on the *ought self* and are better competencies for development.

### **SEM and Competencies**

A direct personality-competency relationship for 360-degree feedback has only been proposed in a limited number of studies (Lin, 2012). Silvester and Wyatt (2012) attempted to utilize CFA to examine different conceptualizations performance only using self-ratings to construct their latent models while ignoring all other types of ratings. The concept of classifying different raters as passive instruments of competency assessment has been proposed by Schmidt and Hunter (1996), but researchers have failed to utilize the similarities of SEM's construct-indicator relationships and 360-degree feedback's competency-item relationships. However, similar to Mount et al. (1998), the present research hypothesized that in 360-degree appraisals, raters from different perspectives are measuring different aspects of performance, and analogous to Schmidt and Hunter's (1996) research, each rater can serve as an item passively measuring performance. Combining these two propositions, the present study suggested that 360-degree feedback competencies can be constructed as latent variables utilizing CFA and SEM, and different raters can serve as separate items measuring a latent construct.

Previous research has shown that agreement between self-ratings and ratings provided by others is lower than agreement of ratings provided by peers and supervisors (Harris & Schaubroeck, 1988), and that ratings provided by different sources are likely to be somewhat inconsistent (Murphy & Cleveland, 1991; Cardy & Dobbins, 1994). Rating

inconsistencies may occur because managers behave differently depending on whether they are interacting with peers, subordinates, customers, or supervisors (London & Smither, 1995), or because different raters observe different behaviors (Cardy & Dobbins, 1994). Researchers should consider inter-rater reliability when deciding if it is appropriate to focus on the average rating of each ratee across the raters (London & Smither, 1995), or using higher-order factors when analyzing the effects of individual-level factors on multi-source feedback (Bryk & Raudenbush, 2001). This research study tested both averaged rater scores by item and by rater, while also using higher-order models of rater scores to test the model for fit and factor loadings.

Personality attributes of all kinds may be associated with behavioral ratings in multi-source feedback and reflected in the correlation between the personality construct and the behavior (Warr & Hoare, 2002). Personality variables have been found to significantly predict job behavior criteria (Warr, 1999), and using SEM, Farh and Dobbins (1989) were able to directly predict the effect of personality (self-esteem) on supervisor rating dimensions. Findings such as these indicate that SEM can be used successfully to assess personality-competency relationships, and this research study tested the personality-competency relationships using SEM.

### **Self-Rating Differences**

Scullen, Mount, and Judge (2003) found that a set of core performance factors are consistent across rater dimensions. Consistency Theory (Korman, 1970) posited that individuals perceive their behavior to be consistent with their self-esteem in order to maintain a consistent self-image. According to Warr (2000), persons with high self-esteem are more likely to over-rate themselves in all respects relative to judgments

made by other people. However, various personality and ability factors influence one's own self-perceptions (Yammarino & Atwater, 1993). Personality-behavior correlations may be larger when behavior ratings are made by the self rather than by a supervisor (Warr, 2000), and self-esteem may affect average self-report scores (Brutus et al., 1999). As previously mentioned, self-evaluations may become more accurate as ratees develop their own schemas related to the performance domains contained in the multi-source instrument (Markus & Sentis, 1982). However, many studies have found a lack of validity in using self-ratings in 360-degree feedback appraisals (Farh et al., 1988; Harris & Schaubroeck, 1988; Warr, 2002). Due to the inconsistencies in results of previous research about self-ratings, the present research examined both including and not including self-ratings in CFA models of 360-degree feedback competencies to test whether self-appraisals are a valid method of assessing competencies as indicated by self-appraisal factor loadings.

### **Hypotheses**

This study used CFA to test the construct validity of hypotheses and SEM to test the relationships between personality and 360-degree feedback competencies. This research was exploratory in nature. Consequently, specific relationships between personality and 360-degree feedback competencies were not hypothesized. Hypotheses were more focused on the methodological effects of using CFA and SEM to compare the two concepts of personality and 360-degree feedback ratings.

### Hypothesis One

Configuring 360-degree feedback ratings into constructs with averaged ratings using a CFA will produce a model with acceptable fit. Even though the model is expected to have suitable fit, the factor loading for self-ratings are expected to be low (i.e.,  $< .50$ ). Figure 3 shows an example of how different 360-degree feedback raters would load on a construct using CFA. The 360-degree feedback competencies that will be tested include: business acumen, driving for results, managing others, planning and organizing, relationship management, resilience, and written communication (Strategic Success Model, 2003).

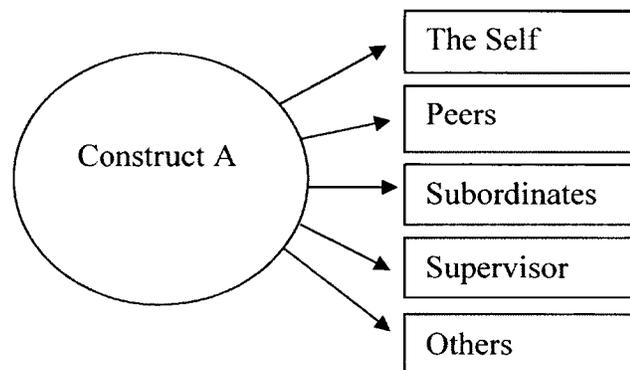


Figure 3 CFA Model for 360-Degree Feedback Competencies

### Hypothesis Two

CFA models of 360-degree feedback competencies with averaged ratings that do not include self-ratings will have better fit than the CFA model including self-ratings (although the fit is not expected to change a considerable amount). Factor loadings for all four of the variables measured are expected to be higher than the self-appraisal factor

loading from Hypothesis One. Figure 4 illustrates the absence of self-ratings from the CFA. All of the 360-degree feedback competencies from Hypothesis One will be tested.

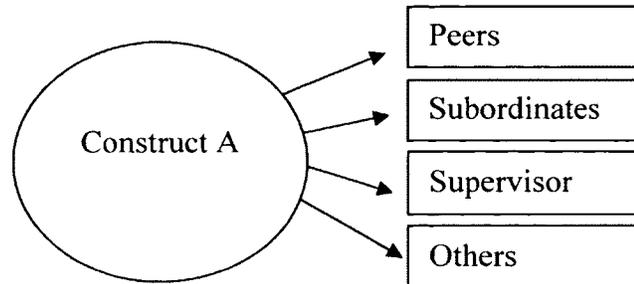


Figure 4 CFA Model with No Self-Ratings

### Hypothesis Three

CFA models of GZTS/DFOS personality inventory variables using item-level ratings will produce a model with acceptable fit. The personality traits measured by the GZTS/DFOS personality inventory that will be used as surrogates of the Big Five include: sociability (extraversion), self-reliance (conscientiousness), need to be liked (agreeableness), emotional evenness (emotional stability), and liking for thinking (openness to experience; Technical Manual, 2009). Figure 5 illustrates the personality variables relationship to the construct using CFA.

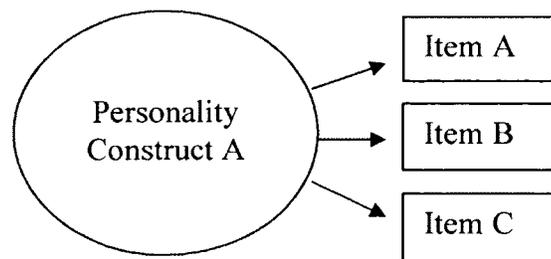


Figure 5 CFA Model for GZTS/DFOS Personality Variables

### Hypothesis Four

Hierarchical models of 360-degree feedback competencies, treating rater types as constructs and superordinate competencies as higher-order factors, will produce models with better fit than using averaged ratings across raters to create first-order factors.

Below, Figure 6 illustrates the use of a higher-order factor in a CFA. Each of the 360-degree feedback competencies from Hypothesis One will be tested as its own CFA model.

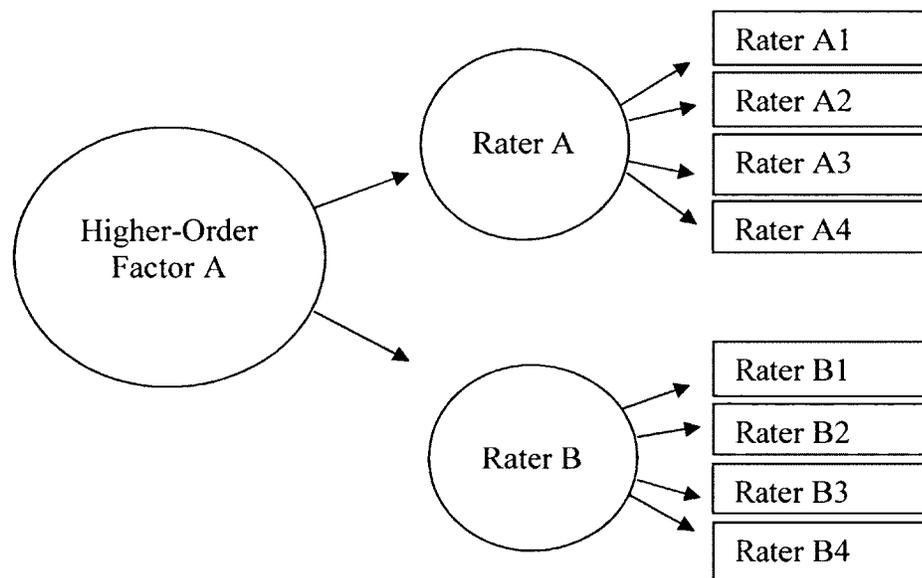


Figure 6 SEM with Higher-Order Competency Factor

### Hypothesis Five

Personality constructs measured by the GZTS/DFOS personality inventory will predict the 360-degree feedback competencies related to traits. A model including personality constructs and trait-based competencies will produce suitable fit. All of the personality variables from Hypothesis Three will be tested, and the model from

Hypotheses One-Four with the best fit indices will be used for the trait-based competencies. The trait-based 360-degree feedback competencies that will be used include: driving for results, managing others, planning and organizing, relationship management, and resilience (Strategic Success Model, 2003). All the personality traits from the GZTS/DFOS personality inventory are expected to be related to trait-based 360-degree feedback competencies. Figure 7 illustrates the personality constructs relationship to trait-based competencies using SEM.

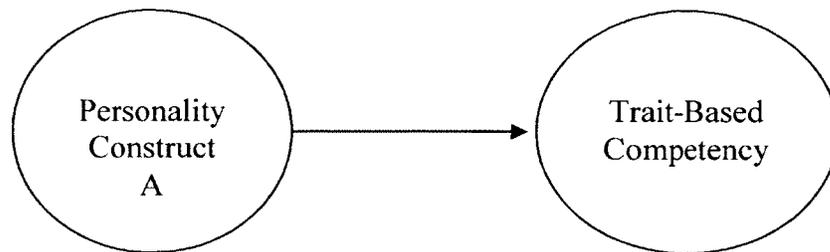


Figure 7 Personality Traits Predicting Trait-Based Competencies

### **Hypothesis Six**

Personality constructs measured by the GZTS/DFOS personality inventory will not predict 360-degree feedback competencies related to skills or behaviors. A model including personality constructs and skill-based competencies will not produce suitable fit. All of the personality variables from Hypothesis Three will be tested, and the model from Hypotheses One through Four with the best fit indices will be used for the skill-based competencies. The skill-based 360-degree feedback competencies that will be used include: business acumen and written communication (Strategic Success Model, 2003). None of the personality traits from the GZTS/DFOS personality inventory are

expected to be related to the skill-based 360-degree feedback competencies. Figure 8 illustrates the personality constructs relationship to skill-based competencies using SEM.

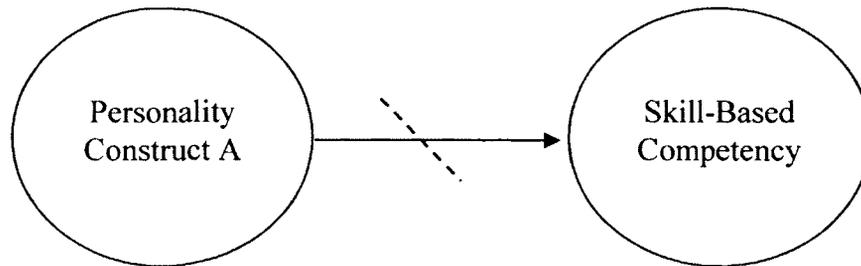


Figure 8 Personality Traits Not Predicting Skill-Based Competencies

## **CHAPTER TWO**

### **METHOD**

#### **Participants**

A dataset including approximately 3,500 participants was provided by a third party consulting firm in the southwestern United States for this research. The sample consisted of middle managers from multiple industries and organizations. The participants included a diverse sample; however, specific demographics were not reported under the non-disclosure agreement. Participants had already received personal feedback on their personality and 360-degree feedback assessments. All identifying information about participants was removed, and six-digit numbers were assigned to each participant by the consulting organization.

#### **Measures**

##### **GZTS/DFOS Personality Inventory**

The GZTS/DFOS personality inventory consists of combining items of the Guilford-Zimmerman Temperament Survey (GZTS) and the Dynamic Factors Opinion Survey (Technical Manual, 2009). The Guilford-Zimmerman personality survey was a widely used personality inventory (Guilford, Zimmerman, and Guilford, 1976), and the DFOS was also developed by Guilford (Guilford & Martin, 1944) with a focus on the assessment of motivational and human needs (e.g., need for attention). All

scales on the GZTS and DFOS have acceptable levels of reliability as evidenced by internal consistency coefficients found in previous research (Technical Manual, 2009).

The GZTS (Guilford & Zimmerman, 1949) consists of 300 items representing 10 personality and temperament factors: general activity (energy vs. inactivity), restraint (seriousness vs. impulsiveness), ascendance (social boldness vs. submissiveness), sociability (social interest vs. shyness), emotional stability (evenness in mood vs. fluctuation of moods), objectivity (thick-skinned vs. hypersensitive), friendliness (agreeableness vs. belligerence), thoughtfulness (reflective vs. disconnected), personal relations (tolerance vs. hypercritical), and masculinity (hardboiled vs. sympathetic). rushton and irwing (2009) subjected the GZTS to a confirmatory factor analysis in which the ten factors were found to have the appropriate fit and validity indices. Guilford and Martin (1944) also derived ten factors from the DFOS. The combined GZTS/DFOS personality inventory also includes two additional faking scales: subtle faking (positive response factor one) and gross faking (positive response factor two; Technical Manual, 2009).

The two personality measures were first combined to make a 480 item personality inventory, which was used for over 10 years before its first re-validation, with subsequent item reduction analyses resulting in the 350 item measure currently used (Technical Manual, 2009). The combined GZTS/DFOS personality inventory is based on professional judgment, substantial research, fair assessment, and job relevance (Technical Manual, 2009). Items on the combined GZTS/DFOS personality inventory were rated on a dichotomous scale (1 = endorsed, 0 = not endorsed).

### **Relation to the Big Five**

Glasgow (1999), reported correlations of the GZTS/DFOS combined inventories with the NEO PI-R (a measure of the Big Five) in a sample of 88 professionals. Many significant relationships were reported (see Table 1) with NEO PI-R and the relationships suggested overlap between the two measures. For the purpose of the present research, the following scales of the GZTS/DFOS personality inventory served as surrogates for the corresponding scales of the Big Five based on Glasgow's (1999) research: (a) sociability (SS) measuring extraversion, (b) self-reliance (SR) measuring conscientiousness, (c) need to be liked (FF) measuring agreeableness, (d) emotional evenness (EE) measuring emotional stability, and (e) liking for thinking (LT) measuring openness to experience.

### **Strategic Success Model 360-Degree Feedback Assessment**

Based on a system of expert judgment, the competency-based 360-degree feedback system that was used in this research was intended to evaluate intellectual abilities and work-oriented personality (Strategic Success Model, 2003). The Strategic Success Model is a measure of many of the innate factors that influence or display competencies (Technical Manual, 2009). The Strategic Success Model was created based on expert judgments over multiple iterations. The model identified 38 competencies grouped into three general areas: thinking, working, and relating (Strategic Success Model, 2003).

It was noted that relationships between specific personality characteristics and particular competencies may exist, but these relationships are not expected to be one-to-one relationships (Technical Manual, 2009). An individual characteristic, or combination of characteristics, can impact multiple competency areas (e.g., Assertiveness can impact

how someone manages others, as well as how they might work together on a team), but not all competencies may be impacted by measurable personality or ability traits (e.g., Safety may not be highly impacted by personality traits; Strategic Success Model, 2003). The present research hypothesized that competencies considered trait-based by the researcher would have a significant relationship to personality constructs, while those competencies that were considered to be more skill or behavior-based by the researcher would not have a significant relationship to personality constructs. The competencies used included: business acumen ( $n = 770$ ), driving for results ( $n = 1023$ ), managing others ( $n = 893$ ), planning and organizing ( $n = 572$ ), relationship management ( $n = 602$ ), resilience ( $n = 317$ ), and written communication ( $n = 177$ ). The Strategic Success Model (2003) competency items were rated on a five-point Likert scale, plus an escape option (0 = *cannot rate*, 1 = *strongly disagree*, 2 = *disagree*, 3 = *neutral*, 4 = *agree*, 5 = *strongly agree*). This research followed Schmidt and Hunter's (1996) approach and treated individual raters as passive instruments within 360-degree feedback appraisals. Also, middle-managers' competencies were grouped by five different sources: self, supervisor, peers (averaged), subordinates (averaged), and others (averaged ratings of customers, previous co-workers, etc.).

### **Trait-Based Competencies**

The present research hypothesized that of the seven 360-degree feedback competencies measured in the Strategic Success Model (2003), the researcher identified five of them as trait-based competencies: driving for results, managing others, planning and organizing, relationship management, and resilience. According to the Strategic Success Model (2003), personal qualities that describe driving for results included:

persistence, overcoming obstacles, drive, high expectations, and an achievement orientation. Qualities that described managing others included: directing, learning, motivating, being fair or objective, being personally accountable, and leadership. Qualities that described planning and organizing included: realism, time management, competence, and consistency. Qualities that described relationship management included: positivity, valuing relationships, sociability, thoughtfulness, and collaboration. The qualities of the Strategic Success Model (2003) that described resilience included: positivity, being even-keeled, lacking stress or frustration, and a recovery orientation.

### **Skill-Based Competencies**

The present research hypothesized that of the seven 360-degree feedback competencies measured in the Strategic Success Model (2003), the researcher identified two of them as skill-based competencies: (a) business acumen and (b) written communication. business acumen is described as understanding business concepts and company's finances, and using knowledge to be an effective manager (Strategic Success Model, 2003). A description of the written communication states the competency is comprised of having the skills to communicate in a written format, articulation of thought, and adjusting writing style to accommodate the audience of the message (Strategic Success Model, 2003).

## **Procedure**

### **Data Screening**

Once the personality and 360-degree feedback data was obtained from the third party consulting firm, the data set was examined for missing data, miscoded items, and assessment of normality assumptions. In cases of missing data and miscoded items, list-wise deletion was employed because the large number of participants (Hair et al., 2006). Missing data was also checked for randomness. If missing data was not systematic, then standard procedures of data screening (e.g., list-wise deletion) were employed. Although approximately 3,500 participants completed the personality measure, only a limited number of the Strategic Success Model competencies were selected by the participating middle-managers. No managers were rated on all 38 competencies of the Strategic Success Model. Consequently, each 360-degree feedback competency was not be rated by all 3,500 participants. Typically, each of the 360-degree feedback competencies had approximately 300-1,200 participants analyzed after the data screening processes were completed.

### **Data Analysis**

The first step of data analysis consisted of using confirmatory factor analysis to assess constructs based on both the GZTS/DFOS personality inventory and the Strategic Success Model 360-degree feedback measure. Seven confirmatory factor analyses (one for each competency) for each hypothesis were performed to assess the measurement and latent model differences using chi-square statistics as well as goodness-of-fit indices, such as the Comparative Fit Index (CFI), and badness-of-fit indices, such as the root-mean-square error of approximation (RMSEA). Each of the seven CFAs was altered to

test Hypotheses One through Four. Traditionally accepted values indicating acceptable fit for CFI are .90 or above and RMSEA values of .07 or if the number of variables is over thirty and  $n > 250$  (Hair et al., 2006). Also, the models were assessed for convergent and discriminant validity (Hair et al., 2006). As mentioned previously and consistent with conventional structural equation analysis (Farh & Dobbins, 1989), Greek letters were used to depict parameters estimated, circles representing latent constructs, and boxes indicating item measures.

A  $\zeta$  (Ksi) indicated an exogenous variable or independent variable construct, while a  $\eta$  (Eta) indicated an endogenous variable or dependent variable construct (Farh & Dobbins, 1989). An example of a CFA model representing Hypothesis One is shown below, and a similar version of this model applies to Hypotheses Two through Four. Seven CFA models were conducted in this research, one for each of the seven 360-degree feedback competencies in Hypotheses One, Two, and Four (Hypothesis Three was subsumed in Hypothesis One's model). In this research, each  $\zeta$  had the appropriate item boxes attached to it with items of the personality or 360-degree feedback measure applicable to the corresponding hypothesis, but due to the large number of boxes necessary to for this CFA, they are not included in the Figure 9.

In this CFA model,  $\zeta_{1-5}$  represented the five personality constructs measured, while  $\zeta_6$  represented the 360-degree feedback competency. The double-headed arrows indicated phi-coefficients (correlations) between the personality constructs and the 360-degree feedback competency. After finding the chi-square and fit indices of the models and hypotheses being tested, the models were transformed into SEM models if all the assumptions of construct validity were met.

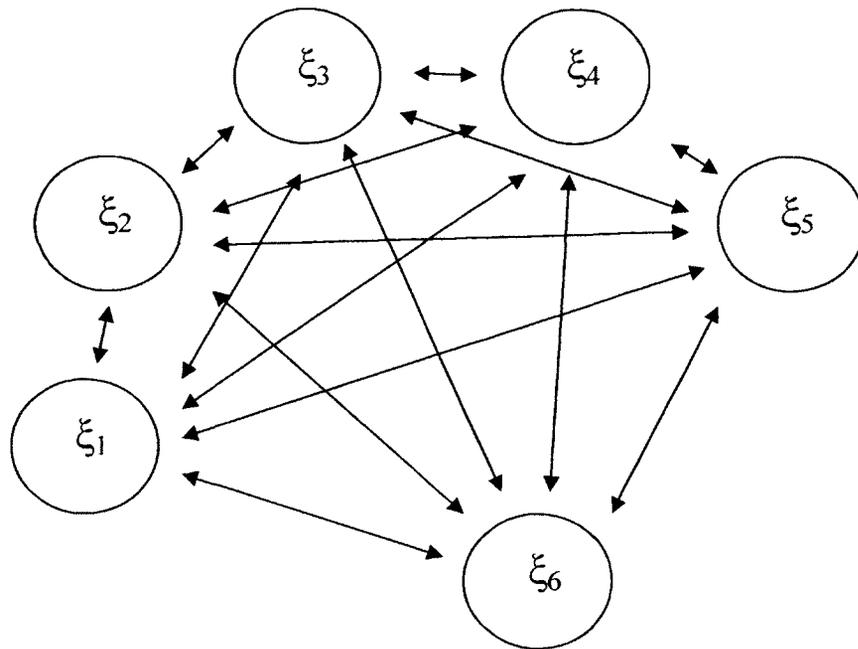


Figure 9 CFA Model for Hypothesis One-Three Without Measured Variable

The model from Hypotheses One through Four with the best fit statistics was converted into an SEM model to test Hypotheses Five and Six. Hypotheses Five and Six were tested with the model shown in Figure 10, and with Hypothesis Four model happening to have the best fit, the model below was altered to indicate a higher-order factor for  $\eta_1$ . Seven SEM models were analyzed in this research, one for each of the seven 360-degree feedback competencies in Hypotheses Five and Six.

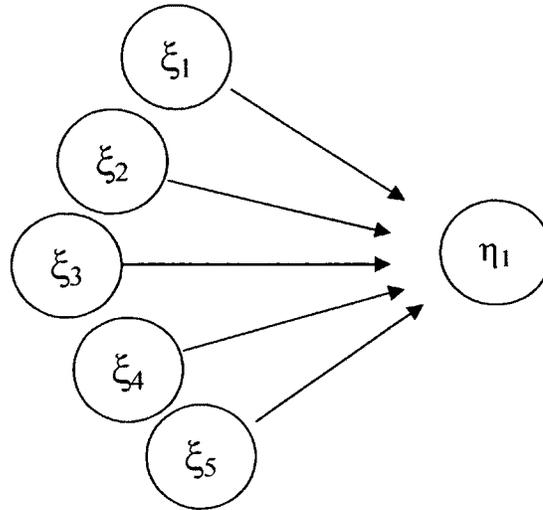


Figure 10 SEM Model for Hypotheses Four-Six Without Measured Variables

In this SEM model,  $\xi_{1-5}$  represented the five personality constructs measured, while  $\eta_1$  represented the 360-degree feedback competency from Hypothesis Four. The model structure was the same regardless of whether trait-based or skill-based competencies were being tested. Although they are missing in the diagram, phi-coefficients (double-headed arrows) indicated correlations between the personality constructs, as is necessary for exogenous variables using SEM path analysis (Bagozzi, 1994). Similar to the CFA models, the SEM model hypotheses were tested by finding the model's chi-square statistic and fit indices. Although not explicitly stated in the hypotheses, some expected personality-competency factor loading relationships included: (a) sociability predicting relationship management and managing others, (b) self-reliance predicting planning and organizing, driving for results, and resilience, (c) need to be liked predicting relationship management and managing others, (d) emotional evenness predicting relationship management and resilience, and (e) liking for thinking predicting driving for results and (negatively) planning and organizing. However, it was noted that

because of the exploratory nature of this research, all personality variables were hypothesized to be related to trait-based competencies, but none were hypothesized to be related to skill-based competencies.

## CHAPTER THREE

### RESULTS

#### Confirmatory Factor Analyses

The first research question (Hypothesis One) proposed there would be evidence of construct validity using 360-degree feedback competency ratings in which ratings were averaged across items by rater type. Seven covariance matrices were constructed to compute the overall validity characteristics of the seven 360-degree feedback competencies. Fit indices (seen in Table 2) indicated significant ( $p < 0.001$ ) and relatively high chi-squares for all seven of the constructs being measured. Traditionally accepted values indicating acceptable fit for the Comparative Fit Index (CFI) are 0.90 or above and 0.07 or below for RMSEA values if there are more than thirty variables and  $n > 250$  (Hair et al., 2006). However, the present researches findings for these two indices were unusual. CFIs ranged from 0.811 to 0.886, which did not meet the acceptable level for goodness-of-fit, while the RMSEAs ranged from 0.047 to 0.060, which were within the acceptable levels for badness-of-fit. The construct that provided the most satisfactory fit evidence was business acumen, with a CFI of 0.89, a RMSEA of 0.047, and chi-square ( $\chi^2$ ) of 901.8 ( $df = 335$ ).

Table 2

## Comparative Model Fit Indices: Hypothesis One

	BA	DfR	MO	P&O	RM	R	WC
$\chi^2$	901.800	1194.000	1004.000	949.700	913.100	683.700	588.400
<i>df</i>	335.000	362.000	362.000	335.000	362.000	362.000	362.000
CFI	0.886	0.871	0.881	0.823	0.860	0.856	0.811
RMSEA	0.047	0.047	0.045	0.057	0.050	0.053	0.060

*Note.* All chi-squares were significant at  $p < 0.001$ . BA = Business Acumen, DfR = Drive for Results, MO = Managing Others, P&O = Planning & Organizing, RM = Relationship Management, R = Resilience, WC = Written Communication.

Construct validity also can be inferred from empirical findings, such as factor loading estimates, construct reliabilities, variance extracted percentages, and inter-construct correlations (Bagozzi & Yi, 2012). Table 3 displays the standardized factor loading estimates for all seven constructs and their items (i.e., self, supervisor, etc.). All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.16 to 0.81. Although factor loadings were fairly consistent across all seven constructs, they were lower for self-ratings. For example, self-ratings ranged from 0.16 to 0.49, while all other types of ratings ranged from 0.46 to 0.81. Furthermore, for the estimated variance extracted to be interpreted as evidence of convergent validity, over 50 percent variance extracted must be estimated from a construct (Hair et al., 2006). By testing Hypothesis One, it was found that the estimated variance extracted ranged from 0.24 to 0.43, none of which met the 50 percent threshold necessary to support the conclusion of convergent validity. For construct reliability estimates to be interpreted as evidence of convergent validity (similar to correlation coefficients), a threshold of over 0.70 must be supported (Hair et al., 2006). The construct reliability estimates ranged from 0.58 to 0.79, with only business acumen and resilience over the necessary threshold. Although the results were

mixed for each of the seven constructs, with business acumen meeting the most criteria, construct validity was not supported for the seven measurement models of Hypothesis One.

Table 3

## Item Content and Standardized Factor Loading Estimates: Hypothesis One

	BA	DfR	MO	P&O	RM	R	WC
BA-Self	0.49						
BA-Super	0.62						
BA-Peer	0.65						
BA-Sub	0.71						
BA-Other	0.77						
DfR-Self		0.34					
DfR-Super		0.54					
DfR-Peer		0.67					
DfR-Sub		0.58					
DfR-Other		0.58					
MO-Self			0.16				
MO-Super			0.47				
MO-Peer			0.52				
MO-Sub			0.60				
MO-Other			0.56				
P&O-Self				0.35			
P&O-Super				0.55			
P&O-Peer				0.59			
P&O-Sub				0.57			
P&O-Other				0.56			
RM-Self					0.33		
RM-Super					0.56		
RM-Peer					0.68		
RM-Sub					0.58		
RM-Other					0.61		
R-Self						0.42	
R-Super						0.53	
R-Peer						0.81	
R-Sub						0.62	
R-Other						0.62	
WC-Self							0.23
WC-Super							0.55
WC-Peer							0.65
WC-Sub							0.46
WC-Other							0.47
Variance Extracted	0.43	0.31	0.24	0.28	0.32	0.38	0.24
Construct Reliability	0.79	0.68	0.58	0.66	0.69	0.74	0.60

*Note.* BA = Business Acumen, DfR = Drive for Results, MO = Managing Others, P&O = Planning & Organizing, RM = Relationship Management, R = Resilience, WC = Written Communication.

The second hypothesis proposed finding evidence of construct validity using 360-degree feedback competency ratings for ratings averaged across items by rater type, while excluding self-ratings. Seven covariance matrices were constructed to compute the overall validity characteristics of the seven 360-degree feedback competencies. Fit indices (seen in Table 4) indicated significant ( $p < 0.001$ ) and relatively high chi-squares for all seven of the constructs being measured, although chi-squares were found to be lower at face value than those found in the Hypothesis One analyses. The fit indices were similar to the previous hypothesis in that CFIs ranged from 0.826 to 0.894, which did not meet an acceptable level for goodness-of-fit, but the RMSEAs ranged from 0.043 to 0.059, all of which were within the acceptable levels for badness-of-fit. The fit indices of Hypothesis Two were more clearly related to fit indices requirements across all seven constructs than the fit indices of Hypothesis One. The construct for which there was the clearest evidence of satisfactory fit evidence was managing others, with a CFI of 0.89, a RMSEA of 0.043, and chi-square of 894.5 ( $df = 335$ ).

Table 4

## Comparative Model Fit Indices: Hypothesis Two

	BA	DfR	MO	P&O	RM	R	WC
$\chi^2$	845.600	1083.500	894.500	868.900	800.100	615.700	539.100
$df$	309.000	335.000	335.000	309.000	335.000	335.000	335.000
CFI	0.888	0.881	0.894	0.834	0.878	0.869	0.826
RMSEA	0.048	0.047	0.043	0.056	0.048	0.051	0.059

*Note.* All chi-squares were significant at  $p < 0.001$ . BA = Business Acumen, DfR = Drive for Results, MO = Managing Others, P&O= Planning & Organizing, RM = Relationship Management, R = Resilience, WC = Written Communication.

Table 5 displays the standardized factor loading estimates for all seven constructs and their items (i.e., self, supervisor, etc.). All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.44 to 0.83. When self-ratings are excluded, factor loading estimates improved for all seven constructs. However, the estimated variance extracted ranged from 0.28 to 0.44, none of which met the 50 percent threshold necessary for convergent validity. The construct reliability estimates ranged from 0.61 to 0.75, with only business acumen, relationship management, and resilience having met the necessary threshold of 0.70. While relationship management did not meet the construct validity threshold in Hypothesis One, it did meet the threshold for Hypothesis Two. Because the results were mixed for all seven constructs, construct validity was not supported for the seven measurement models of Hypothesis Two.

The third research question (Hypothesis Three) proposed finding evidence of construct validity within the personality scales of the GZTS/DFOS personality survey related to the Big Five. All of the personality scales were included in the CFAs conducted on the 360-degree feedback data examined in Hypothesis One and Two (see Tables 2 and 4). Consequently, fit indices are included in this section.

Table 5

## Item Content and Standardized Factor Loading Estimates: Hypothesis Two

	BA	DfR	MO	P&O	RM	R	WC
BA-Super	0.61						
BA-Peer	0.64						
BA-Sub	0.72						
BA-Other	0.66						
DfR-Super		0.55					
DfR-Peer		0.67					
DfR-Sub		0.56					
DfR-Other		0.59					
MO-Super			0.47				
MO-Peer			0.52				
MO-Sub			0.60				
MO-Other			0.56				
P&O-Super				0.51			
P&O-Peer				0.61			
P&O-Sub				0.56			
P&O-Other				0.60			
RM-Super					0.57		
RM-Peer					0.69		
RM-Sub					0.57		
RM-Other					0.61		
R-Super						0.51	
R-Peer						0.83	
R-Sub						0.64	
R-Other						0.60	
WC-Super							0.51
WC-Peer							0.68
WC-Sub							0.44
WC-Other							0.47
Variance Extracted	0.44	0.35	0.29	0.33	0.38	0.43	0.28
Construct Reliability	0.75	0.68	0.62	0.66	0.70	0.74	0.61

*Note.* BA = Business Acumen, DfR = Drive for Results, MO = Managing Others, P&O = Planning & Organizing, RM = Relationship Management, R = Resilience, WC = Written Communication.

Table 6 displays the standardized factor loading estimates for all five personality constructs and their scale items. All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.26 to 0.77. The estimated variance extracted ranged from 0.25 to 0.50, with only sociability (SS) meeting the 50 percent threshold necessary to infer convergent validity. The construct reliability estimates ranged from 0.59 to 0.83, with only sociability (SS), liking for thinking (LT), and emotional evenness (EE) exceeding the threshold of 0.70. Although the results were mixed for the five personality scales, construct validity was supported for the sociability (SS) scale. Thus, construct validity evidence was not present for all of the personality scales.

The fourth research question (Hypothesis Four) proposed finding evidence of construct validity using 360-degree feedback competency ratings as seven higher-order constructs and using rater-type (i.e., supervisor, peer, etc.) as lower-order factors, while not using averaged ratings across raters. Seven covariance matrices were constructed to compute the overall validity characteristics of the seven 360-degree feedback competency higher-order factors. Fit indices (seen in Table 7) indicated significant ( $p < 0.001$ ) and very high chi-squares for all seven of the constructs being measured. The findings for the fit indices were similar to the previous hypotheses in that CFIs ranged from 0.845 to 0.914, with only planning and organizing meeting the acceptable level for goodness-of-fit, but the RMSEAs ranged from 0.042 to 0.062, all of which were within the acceptable levels for badness-of-fit.

Table 6

## Items and Standardized Factor Loading Estimates: Hypothesis Three

	SS	SR	LT	FF	EE
SS13	0.76				
SS46	0.69				
SS79	0.72				
SS1012	0.65				
SS1315	0.70				
SR13		0.59			
SR46		0.53			
SR79		0.38			
SR1012		0.55			
LT13			0.67		
LT46			0.57		
LT79			0.64		
LT1012			0.69		
FF13				0.42	
FF46				0.66	
FF79				0.26	
FF1012				0.39	
FF1315				0.63	
EE13					0.40
EE46					0.58
EE79					0.75
EE1012					0.71
EE1315					0.77
Variance Extracted	0.50	0.27	0.42	0.25	0.43
Construct Reliability	0.83	0.59	0.74	0.60	0.78

*Note.* SS = Sociability, SR = Self-Reliance, LT = Liking for Thinking, FF = Need to be Liked, EE = Emotional Evenness.

Compared to previous hypotheses, using a higher-order factor model demonstrated satisfactory fit for all seven 360-degree feedback constructs. The construct that provided the most satisfactory fit evidence was planning and organizing, with a CFI of 0.91, a RMSEA of 0.042, and chi-square of 2415.2 ( $df = 1205$ ). It was the only construct to meet all of evidence Hair et al. (2006) outlined for supporting construct validity.

Table 7

## Comparative Model Fit Indices: Hypothesis Four

	BA	DfR	MO	P&O	RM	R	WC
$\chi^2$	2718.900	3801.900	4554.900	2415.200	2035.700	1313.800	1101.600
$df$	683.000	1061.000	1465.000	1205.000	883.000	721.000	721.000
CFI	0.845	0.881	0.850	0.914	0.898	0.888	0.885
RMSEA	0.062	0.050	0.049	0.042	0.047	0.051	0.055

*Note.* All chi-squares were significant at  $p < 0.001$ . BA = Business Acumen, DfR = Drive for Results, MO = Managing Others, P&O = Planning & Organizing, RM = Relationship Management, R = Resilience, WC = Written Communication.

### Business Acumen Higher-Order Construct

Table 8 displays the standardized factor loading estimates for the four lower-order constructs and the one higher-order construct, business acumen. All loading estimates were significant ( $p < 0.001$ ), ranging from 0.66 to 0.87. Convergent validity was assessed by examining the variance extracted and the construct reliabilities (Hair et al., 2006). The variance extracted estimate for the business acumen higher-order construct was 0.48, below the 50 percent threshold necessary for convergent validity. The construct reliability estimate of business acumen was 0.79, which met the necessary threshold. The fit evidence and convergent validity evidence were not conclusive in presenting construct

validity evidence of a business acumen higher-order factor. All lower-order constructs supported the convergent validity of the measurement model.

Table 8

Item Content and Standardized Factor Loading Estimates: Business Acumen

	Supervisor	Peer	Sub	Other	HOF
Supervisor1	0.80				
Supervisor2	0.79				
Supervisor3	0.74				
Supervisor4	0.82				
Peer1		0.87			
Peer2		0.84			
Peer3		0.75			
Peer4		0.83			
Sub1			0.85		
Sub2			0.79		
Sub3			0.85		
Sub4			0.86		
Other1				0.87	
Other2				0.78	
Other3				0.80	
Other4				0.83	
SUPERVISOR <sub>v</sub>					0.66
PEER <sub>v</sub>					0.75
SUBORDINATE <sub>v</sub>					0.69
OTHER <sub>v</sub>					0.68
Variance Extracted	0.62	0.68	0.70	0.67	0.48
Construct Reliability	0.87	0.89	0.90	0.89	0.79

*Note.* Sub = Subordinate, HOF = Higher-Order Factor (Business Acumen), v = indicates Lower-Order Construct.

Discriminant validity was also tested for the higher-order business acumen construct. To calculate the discriminant validity of a construct, squared inter-construct correlations (SIC) are calculated and then compared to the average variance extracted (AVE) of the construct. If none of the SICs are greater than the AVEs, discriminant

validity may be supported for the model (Hair et al., 2006). Tables 9 and 10 display the SIC values calculated from the inter-construct correlations as well as the comparison of the SIC values to the AVE values. None of the SICs were greater than the AVEs for business acumen, providing evidence supporting discriminant validity. For the higher-order construct of business acumen, the fit evidence, convergent validity evidence, and discriminant validity evidence presented partial construct validity evidence of the measurement model. For the model to meet the conditions of construct validity, higher CFIs would be needed ( $> 0.90$ ), higher AVEs ( $> 0.50$ ), and non-significant chi-squares.

Table 9

## Discriminant Validity Squared Inter-Construct Correlations: Business Acumen

	IC	SIC
EE --- LT	.039	.002
EE --- SR	.298	.089
EE --- SS	.322	.104
EE --- FF	.443	.196
LT --- SR	.040	.002
LT --- SS	.039	.002
LT --- FF	-.047	.002
SR --- SS	.070	.005
FF --- SR	.192	.037
FF --- SS	.229	.052
BUS A --- LT	-.042	.002
BUS A --- SR	.122	.015
BUS A --- SS	-.047	.002
BUS A --- FF	-.005	.000
BUS A --- EE	.043	.002

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, BUS A = Business Acumen (Higher-Order Factor), IC = Inter-construct Correlations, SIC = Squared Inter-construct Correlation.

Table 10

## Discriminant Validity Comparisons: Business Acumen

	AVE	SIC
BUS A	0.48	.002, .015, .002, .000, .002
LT	0.42	.002, .002, .002, .002, .002
SR	0.27	.089, .002, .005, .037, .015
SS	0.50	.104, .002, .005, .052, .002
FF	0.25	.196, .002, .037, .052, .000
EE	0.43	.089, .002, .104, .196, .002

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, BUS A = Business Acumen (Higher-Order Factor), AVE = Average Variance Extracted, SIC = Squared Inter-construct Correlation.

### Drive for Results Higher-Order Construct

Table 11 displays the standardized factor loading estimates for the four lower-order constructs and the higher-order drive for results construct. All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.58 to 0.89. Convergent validity was assessed by examining the variance extracted and the construct reliabilities. The variance extracted estimate for the drive for results higher-order construct was 0.39, which was lower than the 50 percent threshold necessary for convergent validity. The construct reliability estimate of drive for results was 0.72, which met the necessary threshold. For the higher-order construct of drive for results, the fit evidence and convergent validity evidence yielded inconclusive evidence of construct validity. All the lower-order constructs supported the convergent validity of the measurement model.

Table 11

## Item Content and Standardized Factor Loading Estimates: Drive for Results

	Supervisor	Peer	Sub	Other	HOF
Supervisor1	0.75				
Supervisor2	0.66				
Supervisor3	0.72				
Supervisor4	0.76				
Supervisor5	0.75				
Supervisor6	0.76				
Peer1		0.80			
Peer2		0.85			
Peer3		0.81			
Peer4		0.79			
Peer5		0.73			
Peer6		0.77			
Sub1			0.87		
Sub2			0.89		
Sub3			0.86		
Sub4			0.84		
Sub5			0.76		
Sub6			0.79		
Other1				0.82	
Other2				0.84	
Other3				0.84	
Other4				0.84	
Other5				0.73	
Other6				0.76	
SUPERVISOR <sub>v</sub>					0.58
PEER <sub>v</sub>					0.71
SUBORDINATE <sub>v</sub>					0.58
OTHER <sub>v</sub>					0.62
Variance Extracted	0.54	0.63	0.70	0.65	0.39
Construct Reliability	0.87	0.91	0.93	0.92	0.72

*Note.* Sub = Subordinate, HOF = Higher-Order Factor (Drive for Results), v = indicates Lower-Order Construct.

For construct validity to be supported, discriminant validity must also be shown for the higher-order drive for results construct. The discriminant validity was calculated

by comparing the squared inter-construct correlations (SIC) to the average variance extracted (AVE) of the construct. If none of the SICs were greater than the AVEs, discriminant validity is supported for the model. Tables 12 and 13 display the SIC values calculated from the inter-construct correlations as well as the comparison of the SIC values to the AVE values. None of the SICs were greater than the AVEs for drive for results, providing evidence supporting discriminant validity. For the higher-order construct of drive for results, the fit evidence, convergent validity evidence, and discriminant validity evidence partially supported the construct validity of the measurement model. For the model to meet the conditions of construct validity, higher CFIs would be needed ( $> 0.90$ ), higher AVEs ( $> 0.50$ ), and non-significant chi-squares.

Table 12

## Discriminant Validity Squared Inter-Construct Correlations: Drive for Results

	IC	SIC
SS --- SR	.035	.001
SS --- LT	.030	.001
SS --- FF	.167	.028
SS --- EE	.332	.110
SR --- LT	.010	.000
SR --- FF	.171	.029
SR --- EE	.259	.067
LT --- FF	-.069	.005
LT --- EE	.055	.003
FF --- EE	.427	.182
DRIV --- SS	.068	.005
DRIV --- SR	-.012	.000
DRIV --- LT	-.071	.005
DRIV --- FF	.008	.000
DRIV --- EE	.005	.000

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, DRIV = Drive for Results (Higher-Order Factor), IC = Inter-construct Correlations, SIC = Squared Inter-construct Correlation.

Table 13

## Discriminant Validity Comparisons: Driving for Results

	AVE	SIC
DRIV	0.39	.005, .000, .005, .000, .000
LT	0.42	.001, .000, .005, .003, .005
SR	0.27	.001, .000, .029, .067, .000
SS	0.50	.001, .001, .028, .110, .005
FF	0.25	.028, .029, .005, .182, .000
EE	0.43	.110, .067, .003, .182, .000

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, DRIV = Drive for Results (Higher-Order Factor), AVE = Average Variance Extracted, SIC = Squared Inter-construct Correlation.

### Managing Others Higher-Order Construct

Table 14 displays the standardized factor loading estimates for managing others, the higher-order construct and the four lower-order constructs. All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.51 to 0.82. Convergent validity was assessed by examining the variance extracted and the construct reliabilities. The variance extracted estimate for the managing others higher-order construct was 0.32, which was lower than the 50 percent threshold necessary for convergent validity. The construct reliability estimate of managing others was 0.65, which was also lower than the necessary threshold. For the managing others higher-order construct, the fit evidence and convergent validity evidence were not supportive of construct validity. All of the lower-order constructs supported the convergent validity of the measurement model.

Table 14

## Item Content and Standardized Factor Loading Estimates: Managing Others

	Supervisor	Peer	Sub	Other	HOF
Supervisor1	0.66				
Supervisor2	0.63				
Supervisor3	0.75				
Supervisor4	0.56				
Supervisor5	0.63				
Supervisor6	0.65				
Supervisor7	0.67				
Peer1		0.62			
Peer2		0.72			
Peer3		0.77			
Peer4		0.72			
Peer5		0.66			
Peer6		0.81			
Peer7		0.70			
Sub1			0.71		
Sub2			0.82		
Sub3			0.78		
Sub4			0.81		
Sub5			0.69		
Sub6			0.83		
Sub7			0.67		
Other1				0.60	
Other2				0.72	
Other3				0.78	
Other4				0.80	
Other5				0.67	
Other6				0.81	
Other7				0.69	
SUPERVISOR <sub>v</sub>					0.51
PEER <sub>v</sub>					0.55
SUBORDINATE <sub>v</sub>					0.61
OTHER <sub>v</sub>					0.59
Variance Extracted	0.42	0.52	0.58	0.53	0.32
Construct Reliability	0.85	0.90	0.92	0.90	0.65

*Note.* Sub = Subordinate, HOF = Higher-Order Factor (Managing Others), v = indicates Lower-Order Construct.

Discriminant validity was also tested for the higher-order construct of managing others. The discriminant validity was calculated by comparing the squared inter-construct correlations (SIC) to the average variance extracted (AVE) of the construct. If none of the SICs were greater than the AVEs, discriminant validity would be supported for the model. Tables 15 and 16 display the SIC values calculated from the inter-construct correlations as well as the comparison of the SIC values to the AVE values. None of the SICs were greater than the AVEs for managing others, providing evidence for discriminant validity. For the higher-order construct of managing others, the partial fit evidence, no convergent validity evidence, and discriminant validity evidence provided mixed evidence of construct validity for the measurement model. For the model to meet the conditions of construct validity, higher CFIs would be needed ( $> 0.90$ ), higher AVEs ( $> 0.50$ ), higher construct reliabilities ( $> 0.70$ ), and non-significant chi-squares.

Table 15

## Discriminant Validity Squared Inter-Construct Correlations: Managing Others

	IC	SIC
SR --- SS	.047	.002
LT --- SS	.067	.004
FF --- SS	.143	.020
EE --- SS	.374	.140
LT --- SR	.028	.001
FF --- SR	.272	.074
EE --- SR	.294	.086
FF --- LT	-.103	.011
EE --- LT	.075	.006
EE --- FF	.434	.188
MAN O --- SS	-.018	.000
MAN O --- SR	-.070	.005
MAN O --- LT	-.113	.013
MAN O --- FF	.016	.000
MAN O --- EE	.008	.000

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, MAN O = Managing Others (Higher-Order Factor), IC = Inter-construct Correlations, SIC = Squared Inter-construct Correlation.

Table 16

## Discriminant Validity Comparisons: Managing Others

	AVE	SIC
MAN O	0.32	.000, .005, .013, .000, .000
LT	0.42	.004, .001, .011, .006, .013
SR	0.27	.002, .001, .074, .086, .005
SS	0.50	.002, .004, .020, .140, .000
FF	0.25	.020, .074, .011, .188, .000
EE	0.43	.140, .086, .006, .188, .000

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, MAN O = Managing Others (Higher-Order Factor), AVE = Average Variance Extracted, SIC = Squared Inter-construct Correlation.

### **Planning & Organizing Higher-Order Construct**

Table 17 displays the standardized factor loading estimates for the planning and organizing higher-order construct and the four lower-order constructs. All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.54 to 0.91. Convergent validity was assessed by examining the variance extracted and the construct reliabilities. The variance extracted estimate for the planning and organizing higher-order construct was 0.35, which was lower than the 50 percent threshold necessary for convergent validity. The construct reliability estimate of planning and organizing was 0.68, which was lower than the necessary threshold. For the planning and organizing higher-order construct, the fit evidence and convergent validity evidence were not conclusive in supporting the construct validity of a higher-order factor. All the lower-order constructs supported the convergent validity of the measurement model.

Table 17

## Item Content and Standardized Factor Loading Estimates: Planning &amp; Organizing

	Supervisor	Peer	Sub	Other	HOF
Supervisor1	0.89				
Supervisor2	0.76				
Supervisor3	0.68				
Supervisor4	0.74				
Supervisor5	0.64				
Supervisor6	0.79				
Supervisor7	0.76				
Peer1		0.73			
Peer2		0.81			
Peer3		0.72			
Peer4		0.77			
Peer5		0.72			
Peer6		0.79			
Peer7		0.89			
Sub1			0.85		
Sub2			0.86		
Sub3			0.81		
Sub4			0.82		
Sub5			0.76		
Sub6			0.82		
Sub7			0.91		
Other1				0.77	
Other2				0.82	
Other3				0.73	
Other4				0.76	
Other5				0.73	
Other6				0.77	
Other7				0.90	
SUPERVISOR <sub>v</sub>					0.54
PEER <sub>v</sub>					0.64
SUBORDINATE <sub>v</sub>					0.57
OTHER <sub>v</sub>					0.62
Variance Extracted	0.57	0.60	0.70	0.62	0.35
Construct Reliability	0.90	0.91	0.94	0.92	0.68

*Note.* Sub = Subordinate, HOF = Higher-Order Factor (Planning & Organizing), v = indicates Lower-Order Construct.

Discriminant validity was also tested for the higher-order construct of planning and organizing. The discriminant validity was calculated by comparing the squared inter-construct correlations (SIC) to the variance extracted (AVE) of the construct. If none of the SICs were greater than the AVEs, discriminant validity was supported for the model. Tables 18 and 19 display the SIC values calculated from the inter-construct correlations as well as the comparison of the SIC values to the AVE values. None of the SICs were greater than the AVEs for planning and organizing, providing evidence supporting discriminant validity.

For the higher-order construct of planning and organizing, the fit evidence, convergent validity evidence, and discriminant validity evidence partially supported the construct validity of the measurement model. For the model to meet the conditions of construct validity, higher AVEs would be needed ( $> 0.50$ ), higher construct reliabilities ( $> 0.70$ ), and non-significant chi-squares.

Table 18

## Discriminant Validity Squared Inter-Construct Correlations: Planning &amp; Organizing

	IC	SIC
SR --- EE	.310	.096
SR --- SS	.038	.001
EE --- SS	.305	.093
EE --- FF	.457	.209
SS --- FF	.106	.011
SR --- FF	.229	.052
SR --- LT	-.010	.000
SS --- LT	.036	.001
EE --- LT	.025	.001
FF --- LT	-.197	.039
PLAN --- SS	-.023	.001
PLAN --- SR	.011	.000
PLAN --- LT	-.004	.000
PLAN --- FF	.032	.001
PLAN --- EE	.024	.001

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, PLAN = Planning & Organizing (Higher-Order Factor), IC = Inter-construct Correlations, SIC = Squared Inter-construct Correlation.

Table 19

## Discriminant Validity Comparisons: Planning &amp; Organizing

	AVE	SIC
PLAN	0.35	.001, .000, .000, .001, .001
LT	0.42	.000, .001, .001, .039, .000
SR	0.27	.096, .001, .052, .000, .000
SS	0.50	.001, .093, .011, .001, .001
FF	0.25	.209, .011, .052, .039, .001
EE	0.43	.096, .093, .209, .001, .001

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, PLAN = Planning & Organizing (Higher-Order Factor), AVE = Average Variance Extracted, SIC = Squared Inter-construct Correlation.

### **Relationship Management Higher-Order Construct**

Table 20 displays the standardized factor loading estimates for the relationship management higher-order construct and the four lower-order constructs. All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.60 to 0.87. Convergent validity was assessed by examining the variance extracted and the construct reliabilities. The variance extracted estimate for the relationship management higher-order construct was 0.42, which was less than the 50 percent threshold necessary for convergent validity. The construct reliability estimate of relationship management was 0.74, which was above the necessary threshold. For the relationship management higher-order construct, the fit evidence and convergent validity evidence were not conclusive in supporting the construct validity of a higher-order factor. All of the lower-order constructs supported the convergent validity of the measurement model.

Table 20

## Item Content and Standardized Factor Loading Estimates: Relationship Management

	Supervisor	Peer	Sub	Other	HOF
Supervisor1	0.67				
Supervisor2	0.75				
Supervisor3	0.76				
Supervisor4	0.70				
Supervisor5	0.73				
Peer1		0.80			
Peer2		0.82			
Peer3		0.82			
Peer4		0.74			
Peer5		0.81			
Sub1			0.86		
Sub2			0.86		
Sub3			0.85		
Sub4			0.82		
Sub5			0.83		
Other1				0.79	
Other2				0.82	
Other3				0.87	
Other4				0.81	
Other5				0.82	
SUPERVISOR <sub>v</sub>					0.62
PEER <sub>v</sub>					0.72
SUBORDINATE <sub>v</sub>					0.60
OTHER <sub>v</sub>					0.63
Variance Extracted	0.52	0.64	0.71	0.68	0.42
Construct Reliability	0.85	0.90	0.93	0.91	0.74

*Note.* Sub = Subordinate, HOF = Higher-Order Factor (Relationship Management), v = indicates Lower-Order Construct.

To further examine construct validity, discriminant validity was also examined for the higher-order construct of relationship management. This was accomplished by comparing the squared inter-construct correlations (SIC) to the variance extracted (AVE) of the construct. Tables 21 and 22 display the SIC values calculated from the

inter-construct correlations as well as the comparison of the SIC values to the AVE values. None of the SICs were greater than the AVEs for relationship management, providing evidence supporting discriminant validity. For the higher-order construct of relationship management, the fit evidence, convergent validity evidence, and discriminant validity evidence partially supported the evidence of the construct validity for the measurement model. For the model to meet the conditions of construct validity, higher CFIs would be needed ( $> 0.90$ ), higher AVEs ( $> 0.50$ ), and non-significant chi-squares.

Table 21

## Discriminant Validity Squared Inter-Construct Correlations: Relationship Management

	IC	SIC
SS --- SR	.033	.001
SS --- LT	.114	.013
SS --- FF	.127	.016
SS --- EE	.408	.166
SR --- LT	-.042	.002
FF --- SR	.293	.086
SR --- EE	.425	.181
FF --- LT	-.153	.023
EE --- LT	.028	.001
FF --- EE	.402	.162
REL M --- SS	.182	.033
REL M --- SR	.101	.010
REL M --- LT	-.185	.034
REL M --- FF	.125	.016
REL M --- EE	.111	.012

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, REL M = Relationship Management (Higher-Order Factor), IC = Inter-construct Correlations, SIC = Squared Inter-construct Correlation.

Table 22

## Discriminant Validity Comparisons: Relationship Management

	AVE	SIC
REL M	0.42	.033, .010, .034, .016, .012
LT	0.42	.013, .002, .023, .001, .034
SR	0.27	.001, .002, .181, .086, .010
SS	0.50	.001, .013, .016, .166, .033
FF	0.25	.016, .086, .023, .162, .016
EE	0.43	.166, .181, .001, .162, .012

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, REL M = Relationship Management (Higher-Order Factor), AVE = Average Variance Extracted, SIC = Squared Inter-construct Correlation.

**Resilience Higher-Order Construct**

Table 23 displays the standardized factor loading estimates for the resilience higher-order construct and the four lower-order constructs. All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.56 to 0.88. Convergent validity was assessed by examining the variance extracted and the construct reliabilities. The variance extracted estimate for the resilience higher-order construct was 0.49, which was less than the 50 percent threshold necessary for convergent validity. The construct reliability estimate of resilience was 0.79, which was above the necessary threshold. For the resilience higher-order construct, the fit evidence and convergent validity evidence were not conclusive in supporting the construct validity of a higher-order factor. All of the lower-order constructs supported the convergent validity of the measurement model.

Table 23

## Item Content and Standardized Factor Loading Estimates: Resilience

	Supervisor	Peer	Sub	Other	HOF
Supervisor1	0.63				
Supervisor2	0.79				
Supervisor3	0.77				
Supervisor4	0.82				
Peer1		0.84			
Peer2		0.82			
Peer3		0.85			
Peer4		0.69			
Sub1			0.85		
Sub2			0.87		
Sub3			0.89		
Sub4			0.85		
Other1				0.74	
Other2				0.88	
Other3				0.85	
Other4				0.84	
SUPERVISOR <sub>v</sub>					0.56
PEER <sub>v</sub>					0.88
SUBORDINATE <sub>v</sub>					0.66
OTHER <sub>v</sub>					0.65
Variance Extracted	0.57	0.64	0.75	0.69	0.49
Construct Reliability	0.84	0.88	0.92	0.90	0.79

*Note.* Sub = Subordinate, HOF = Higher-Order Factor (Resilience), v = indicates Lower-Order Construct.

Discriminant validity was also tested for the higher-order construct of resilience. The discriminant validity was calculated by comparing the squared inter-construct correlations (SIC) to the variance extracted (AVE) of the construct. If none of the SICs were greater than the AVEs, discriminant validity was supported for the model (Hair et al., 2006). Tables 24 and 25 display the SIC values calculated from the inter-construct correlations as well as the comparison of the SIC values to the AVE values. None of the SICs were greater than the AVEs for resilience, providing evidence supporting

discriminant validity. For the higher-order construct of resilience, the fit evidence, convergent validity evidence, and discriminant validity evidence partially supported the construct validity for the measurement model. For the model to meet the conditions of construct validity, higher CFIs would be needed ( $> 0.90$ ), higher AVEs ( $> 0.50$ ), and non-significant chi-squares.

Table 24

## Discriminant Validity Squared Inter-Construct Correlations: Resilience

	IC	SIC
SS --- SR	.038	.001
SS --- LT	.142	.020
SS --- FF	.096	.009
SS --- EE	.406	.165
SR --- LT	-.094	.009
FF --- SR	.266	.071
SR --- EE	.385	.148
FF --- LT	-.143	.020
EE --- LT	.255	.065
FF --- EE	.311	.097
RES --- SS	.159	.025
RES --- SR	.033	.001
RES --- LT	-.019	.000
RES --- FF	.160	.026
RES --- EE	.248	.062

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, RES = Resilience (Higher-Order Factor), IC = Inter-construct Correlations, SIC = Squared Inter-construct Correlation.

Table 25

## Discriminant Validity Comparisons: Resilience

	AVE	SIC
RES	0.49	.025, .001, .000, .026, .062
LT	0.42	.020, .009, .020, .065, .000
SR	0.27	.001, .009, .071, .148, .000
SS	0.50	.001, .020, .009, .165, .025
FF	0.25	.009, .071, .020, .097, .026
EE	0.43	.165, .148, .065, .097, .062

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, RES = Resilience (Higher-Order Factor), AVE = Average Variance Extracted, SIC = Squared Inter-construct Correlation.

### Written Communication Higher-Order Construct

Table 26 displays the standardized factor loading estimates for the written communication higher-order construct and the four lower-order constructs. All loading estimates were significant ( $p < 0.001$ ) and ranged from 0.46 to 0.94. Convergent validity was assessed by examining the variance extracted and the construct reliabilities. The variance extracted estimate for the written communication higher-order construct was 0.31, which was lower than the 50 percent threshold necessary for convergent validity. The construct reliability estimate of written communication was 0.64, which was also lower than the necessary threshold. For the written communication higher-order construct, the fit evidence and convergent validity evidence did not support the construct validity of a higher-order factor. All the lower-order constructs supported the convergent validity of the measurement model.

Table 26

## Item Content and Standardized Factor Loading Estimates: Written Communication

	Supervisor	Peer	Sub	Other	HOF
Supervisor1	0.79				
Supervisor2	0.88				
Supervisor3	0.78				
Supervisor4	0.84				
Peer1		0.91			
Peer2		0.94			
Peer3		0.88			
Peer4		0.74			
Sub1			0.89		
Sub2			0.95		
Sub3			0.83		
Sub4			0.76		
Other1				0.86	
Other2				0.93	
Other3				0.80	
Other4				0.74	
SUPERVISOR <sub>v</sub>					0.57
PEER <sub>v</sub>					0.68
SUBORDINATE <sub>v</sub>					0.46
OTHER <sub>v</sub>					0.50
Variance Extracted	0.68	0.76	0.74	0.70	0.31
Construct Reliability	0.89	0.93	0.92	0.90	0.64

*Note.* Sub = Subordinate, HOF = Higher-Order Factor (Written Communication), v = indicates Lower-Order Construct.

In an effort to further examine the construct validity of the measurement model, discriminant validity was also examined for the higher-order construct of written communication. Discriminant validity was calculated by comparing the squared inter-construct correlations (SIC) to the variance extracted (AVE) of the construct. Tables 27 and 28 display the SIC values calculated from the inter-construct correlations as well as the comparison of the SIC values to the AVE values. None of the SICs were greater than the AVEs for written communication, providing evidence supporting discriminant

validity. For the higher-order construct of written communication, the fit evidence, the lack of convergent validity evidence, and discriminant validity evidence provided mixed evidence of construct validity for the measurement model. For the model to meet the conditions of construct validity, higher CFIs would be needed ( $> 0.90$ ), higher AVEs ( $> 0.50$ ), higher construct reliabilities ( $> 0.70$ ), and non-significant chi-squares.

Table 27

## Discriminant Validity Squared Inter-Construct Correlations: Written Communication

	IC	SIC
SS --- SR	.088	.008
SS --- LT	.063	.004
SS --- FF	.119	.014
SS --- EE	.305	.093
SR --- LT	-.011	.000
SR --- FF	.290	.084
SR --- EE	.373	.139
LT --- FF	-.006	.000
LT --- EE	-.080	.006
FF --- EE	.456	.208
WRIT --- SS	-.188	.035
WRIT --- SR	.017	.000
WRIT --- LT	-.269	.072
WRIT --- FF	.208	.043
WRIT --- EE	-.187	.035

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, WRIT = Written Communication (Higher-Order Factor), IC = Inter-construct Correlations, SIC = Squared Inter-construct Correlation.

Table 28

## Discriminant Validity Comparisons: Written Communication

	AVE	SIC
WRIT	0.31	.035, .000, .072, .043, .035
LT	0.42	.004, .000, .000, .006, .072
SR	0.27	.008, .000, .084, .139, .000
SS	0.50	.008, .004, .014, .093, .035
FF	0.25	.014, .084, .000, .208, .043
EE	0.43	.093, .139, .006, .208, .035

*Note.* EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, WRIT = Written Communication (Higher-Order Factor), AVE = Average Variance Extracted, SIC = Squared Inter-construct Correlation.

To summarize the fit indices and evidence of convergent validity for all the constructs found in Hypotheses One, Two, and Four, Table 29 was constructed. Several general themes within each hypothesis were noted. First, evidence for one construct being superior (or inferior) to the other constructs across all three hypotheses was not found. business acumen had the best fit for Hypothesis One, managing others had the best fit for Hypothesis Two, and planning and organizing had the best fit for Hypothesis Three. This finding was important because it indicated that all seven constructs could be used as similar measures of performance in future analyses (e.g., only using one 360-degree feedback construct in future analyses, instead of multiple constructs), because no constructs being clearly superior (or inferior) to the other constructs.

However, none of the constructs provided complete evidence of construct validity (i.e., fit, convergent, and discriminant validity). Second, the constructs within the latter hypothesis (Hypothesis Four) had markedly more clear results than the constructs within the previous hypotheses (Hypotheses One and Two) on specific fit indices and convergent validity evidence. The constructs within Hypothesis Two had the lowest

chi-square values across all seven constructs, while the constructs within Hypothesis Three had the most degrees of freedom, highest variance extracted, and highest construct reliabilities. Hypothesis One had the least number of superior fit indices and convergent validity evidence based on the CFAs. The constructs within Hypothesis Three were found to have the most construct validity evidence (with planning and organizing having the most evidence); consequently, the higher-order factor models were transformed into structural models (Hair et al., 2006) for the later analyses.

Table 29

## Comparative Model Fit Indices: CFA Hypotheses Comparison

	BA	DfR	MO	P&O	RM	R	WC
Hyp One							
$\chi^2$	901.8	1194.0	1004.0	949.7	913.1	683.7	588.4
<i>df</i>	335	362	362	335	362	362	362
CFI	0.886	0.871	0.881	0.823	0.860	0.856	0.811
RMSEA	<b>0.047</b>	<b>0.047</b>	0.045	0.057	0.050	0.053	0.060
AVE	0.43	0.31	0.24	0.28	0.32	0.38	0.24
CR	<b>0.79</b>	0.68	0.58	0.66	0.69	0.74	0.60
Hyp Two							
$\chi^2$	<b>845.6</b>	<b>1083.5</b>	<b>894.5</b>	<b>868.9</b>	<b>800.1</b>	<b>615.7</b>	<b>539.1</b>
<i>df</i>	309	335	335	309	335	335	335
CFI	<b>0.888</b>	<b>0.881</b>	<b>0.894</b>	0.834	0.878	0.869	0.826
RMSEA	0.048	<b>0.047</b>	<b>0.043</b>	0.056	0.048	<b>0.051</b>	0.059
AVE	0.44	0.35	0.29	0.33	0.38	0.43	0.28
CR	0.75	0.68	0.62	0.66	0.70	0.74	0.61
Hyp Four							
$\chi^2$	2718.9	3801.9	4554.9	2415.2	2035.7	1313.8	1101.6
<i>df</i>	<b>683</b>	<b>1061</b>	<b>1465</b>	<b>1205</b>	<b>883</b>	<b>721</b>	<b>721</b>
CFI	0.845	<b>0.881</b>	0.850	<b>0.914</b>	<b>0.898</b>	<b>0.888</b>	<b>0.885</b>
RMSEA	0.062	0.050	0.049	<b>0.042</b>	<b>0.047</b>	<b>0.051</b>	<b>0.055</b>
AVE	<b>0.48</b>	<b>0.39</b>	<b>0.32</b>	<b>0.35</b>	<b>0.42</b>	<b>0.49</b>	<b>0.31</b>
CR	<b>0.79</b>	<b>0.72</b>	<b>0.65</b>	<b>0.68</b>	<b>0.74</b>	<b>0.79</b>	<b>0.64</b>

*Note.* All chi-squares were significant at  $p < 0.001$ . Hyp = Hypothesis, BA = Business Acumen, DfR = Drive for Results, MO = Managing Others, P&O = Planning & Organizing, RM = Relationship Management, R = Resilience, WC = Written Communication, AVE = Average Variance Extracted, CR = Construct Reliability, *bolded items are seen as superior for a specific construct.*

## Structural Equation Models

### Trait-Based Structural Models

Table 30 presents the path estimates and the corresponding detail on the structural model for the trait-based 360-degree feedback competency, drive for results. Results indicated that evidence displayed by the model supports the acceptability of the model ( $\chi^2 = 3801.9$ ,  $df = 1061$ ,  $p < 0.001$ , CFI = 0.881, RMSEA = 0.050). In addition, none of the focal path coefficients were significant for personality relationships predicting the drive for results higher-order factor. Thus, Hypothesis Five (i.e., personality significantly predicting trait-based higher-order factors) was not supported for drive for results. Also, the expected relationships in which self-reliance and liking for thinking predict drive for results were not statistically significant. However, evidence of the drive for results higher-order factor structure was supported. Substantial focal path coefficients ( $\beta > 0.500$ ) in which the higher-order factor (drive for results) significantly predicted the lower-order factors (supervisor, peer, subordinate, and other). The drive for results higher-order factor accounted for one percent of the variance in the model.

Table 30

## Structural Model Results: Drive for Results

Relationships	Hypothesized Model
LT → DRIV	n.s.
SR → DRIV	n.s.
SS → DRIV	n.s.
FF → DRIV	n.s.
EE → DRIV	n.s.
DRIV → SUPERVISOR	0.580
DRIV → PEER	0.705
DRIV → SUBORDINATE	0.579
DRIV → OTHER	0.616
Model Fit Indices	
$\chi^2$	3801.9
<i>df</i>	1061
CFI	0.881
RMSEA	0.050
Squared Multiple Correlation (DRIV HOF)	0.010

*Note.* All chi-squares were significant at  $p < 0.001$ . n.s. = not significant, EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, DRIV = Drive for Results (HOF = Higher-Order Factor).

Table 31 presents the path estimates and the corresponding detail on the structural model for the trait-based 360-degree feedback competency, managing others. Results indicated that evidence displayed by the model supports the acceptability of the model ( $\chi^2 = 4554.9$ ,  $df = 1465$ ,  $p < 0.001$ , CFI = 0.850, RMSEA = 0.049). In addition, one of the focal path coefficients (liking for thinking;  $\beta = -0.111$ ) was significant for personality relationships predicting the managing others higher-order factor. Thus, Hypothesis Five (i.e., personality significantly predicting trait-based higher-order factors) was supported for managing others. Also, the expected relationships in which sociability and need to be liked predict managing others were not statistically significant. However, evidence of the managing others higher-order factor structure was supported. Substantial focal path coefficients ( $\beta > 0.500$ ) in which the higher-order factor (managing others) significantly

predicted the lower-order factors (supervisor, peer, subordinate, and other). The managing others higher-order factor accounted for two percent of the variance in the model.

Table 31

## Structural Model Results: Managing Others

Relationships	Hypothesized Model
LT → MAN O	-0.111
SR → MAN O	n.s.
SS → MAN O	n.s.
FF → MAN O	n.s.
EE → MAN O	n.s.
MAN O → SUPERVISOR	0.513
MAN O → PEER	0.554
MAN O → SUBORDINATE	0.606
MAN O → OTHER	0.593
Model Fit Indices	
$\chi^2$	4554.9
<i>df</i>	1465
CFI	0.850
RMSEA	0.049
Squared Multiple Correlation (MAN O HOF)	0.019

*Note.* All chi-squares were significant at  $p < 0.001$ . n.s. = not significant, EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, MAN O = Managing Others (HOF = Higher-Order Factor).

Table 32 presents the path estimates and the corresponding detail on the structural model for the trait-based 360-degree feedback competency, planning and organizing. Results indicated that evidence displayed by the model supports the acceptability of the model ( $\chi^2 = 2415.2$ ,  $df = 1205$ ,  $p < 0.001$ , CFI = 0.914, RMSEA = 0.042). In addition, none of the focal path coefficients were significant for personality relationships predicting the planning and organizing higher-order factor. Thus, Hypothesis Five (i.e., personality predicting trait-based higher-order factors) was not supported for planning &

organizing. Also, the expected relationships in which self-reliance and liking for thinking predict planning and organizing were not statistically significant. However, evidence of the planning and organizing higher-order factor structure was supported. Substantial focal path coefficients ( $\beta > 0.500$ ) in which the higher-order factor (planning & organizing) significantly predicted the lower-order factors (supervisor, peer, subordinate, and other). The planning and organizing higher-order factor accounted for less than one percent of the variance in the model.

Table 32

## Structural Model Results: Planning &amp; Organizing

Relationships	Hypothesized Model
LT → PLAN	n.s.
SR → PLAN	n.s.
SS → PLAN	n.s.
FF → PLAN	n.s.
EE → PLAN	n.s.
PLAN → SUPERVISOR	0.538
PLAN → PEER	0.641
PLAN → SUBORDINATE	0.572
PLAN → OTHER	0.623
Model Fit Indices	
$\chi^2$	2415.2
<i>df</i>	1205
CFI	0.914
RMSEA	0.042
Squared Multiple Correlation (PLAN HOF)	0.002

*Note.* All chi-squares were significant at  $p < 0.001$ . n.s. = not significant, EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, PLAN = Planning & Organizing (HOF = Higher-Order Factor).

Table 33 presents the path estimates and the corresponding detail on the structural model for the trait-based 360-degree feedback competency, relationship management.

Results indicated that evidence displayed by the model supports the acceptability of the

model ( $\chi^2 = 2035.7$ ,  $df = 883$ ,  $p < 0.001$ , CFI = 0.898, RMSEA = 0.047). In addition, two of the focal path coefficients (liking for thinking,  $\beta = -0.196$ ; sociability,  $\beta = 0.204$ ) were significant for personality relationships predicting the relationship management higher-order factor. Thus, Hypothesis Five (i.e., personality significantly predicting trait-based higher-order factors) was supported for relationship management. Also, the expected relationship in which sociability predicts relationship management was supported, while emotional evenness and need to be liked relationships did not predict relationship management. Moreover, evidence of the relationship management higher-order factor structure was supported. Substantial focal path coefficients ( $\beta > 0.500$ ) in which the higher-order factor (relationship management) significantly predicted the lower-order factors (supervisor, peer, subordinate, and other). The relationship management higher-order factor accounted for nine percent of the variance in the model.

Table 33

## Structural Model Results: Relationship Management

Relationships	Hypothesized Model
LT → REL M	-0.196
SR → REL M	n.s.
SS → REL M	0.204 ( <i>confirmed</i> )
FF → REL M	n.s.
EE → REL M	n.s.
REL M → SUPERVISOR	0.618
REL M → PEER	0.722
REL M → SUBORDINATE	0.600
REL M → OTHER	0.634
Model Fit Indices	
$\chi^2$	2035.7
<i>df</i>	883
CFI	0.898
RMSEA	0.047
Squared Multiple Correlation (REL M HOF)	0.086

*Note.* All chi-squares were significant at  $p < 0.001$ . n.s. = not significant, EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, REL M = Relationship Management (HOF = Higher-Order Factor).

Table 34 presents the path estimates and the corresponding detail on the structural model for the trait-based 360-degree feedback competency, resilience. Results indicated that evidence displayed by the model supports the acceptability of the model ( $\chi^2 = 1313.8$ ,  $df = 721$ ,  $p < 0.001$ , CFI = 0.888, RMSEA = 0.051). In addition, one of the focal path coefficients (emotional evenness;  $\beta = 0.256$ ) was significant for personality relationships predicting the resilience higher-order factor. Thus, Hypothesis Five (i.e., personality significantly predicting trait-based higher-order factors) was supported for resilience. Also, the expected relationship in which emotional evenness predicts resilience was supported, while self-reliance did not predict emotional evenness. Moreover, evidence of the resilience higher-order factor structure was supported. Substantial focal path coefficients ( $\beta > 0.500$ ) in which the higher-order factor

(resilience) significantly predicted the lower-order factors (supervisor, peer, subordinate, and other). The resilience higher-order factor accounted for eight percent of the variance in the model.

Table 34

## Structural Model Results: Resilience

Relationships	Hypothesized Model
LT → RES	n.s.
SR → RES	n.s.
SS → RES	n.s.
FF → RES	n.s.
EE → RES	0.256 ( <i>confirmed</i> )
RES → SUPERVISOR	0.556
RES → PEER	0.878
RES → SUBORDINATE	0.659
RES → OTHER	0.645
Model Fit Indices	
$\chi^2$	1313.8
<i>df</i>	721
CFI	0.888
RMSEA	0.051
Squared Multiple Correlation (RES HOF)	0.078

*Note.* All chi-squares were significant at  $p < 0.001$ . n.s. = not significant, EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, RES = Resilience (HOF = Higher-Order Factor).

**Skill-Based Structural Models**

Table 35 presents the path estimates and the corresponding detail on the structural model for the skill-based 360-degree feedback competency, business acumen. Results indicated that evidence displayed by the model supports the acceptability of the model ( $\chi^2 = 2718.9$ ,  $df = 683$ ,  $p < 0.001$ , CFI = 0.845, RMSEA = 0.062). In addition, one of the focal path coefficients (self-reliance;  $\beta = 0.122$ ) was significant in which personality relationships predicted the business acumen higher-order factor. Thus, Hypothesis Six

(i.e., personality will not significantly predict skill-based higher-order factors) was not supported for business acumen. No expected specific personality relationships were hypothesized for skill-based constructs. However, evidence of the business acumen higher-order factor structure was supported. Substantial focal path coefficients ( $\beta > 0.500$ ) in which the higher-order factor (business acumen) significantly predicted the lower-order factors (supervisor, peer, subordinate, and other). The business acumen higher-order factor accounted for two percent of the variance in the model.

Table 35

## Structural Model Results: Business Acumen

Relationships	Hypothesized Model
LT → BUS A	n.s.
SR → BUS A	0.122
SS → BUS A	n.s.
FF → BUS A	n.s.
EE → BUS A	n.s.
BUS A → SUPERVISOR	0.658
BUS A → PEER	0.750
BUS A → SUBORDINATE	0.680
BUS A → OTHER	0.687
Model Fit Indices	
$\chi^2$	2718.9
<i>df</i>	683
CFI	0.845
RMSEA	0.062
Squared Multiple Correlation (BUS A HOF)	0.022

*Note.* All chi-squares were significant at  $p < 0.001$ . n.s. = not significant, EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, BUS A = Business Acumen (HOF = Higher-Order Factor).

Table 36 presents the path estimates and the corresponding detail on the structural model for the skill-based 360-degree feedback competency, written communication.

Results indicated that evidence displayed by the model supports the acceptability of the

model ( $\chi^2 = 1101.6$ ,  $df = 721$ ,  $p < 0.001$ , CFI = 0.885, RMSEA = 0.055). In addition, three of the focal path coefficients (liking for thinking,  $\beta = -0.289$ ; need to be liked,  $\beta = 0.371$ ; emotional evenness,  $\beta = -0.366$ ) were significant in which personality relationships predicted the written communication higher-order factor. Thus, Hypothesis Six (i.e., personality will not significantly predict skill-based higher-order factors) was not supported for written communication. No expected specific personality relationships were hypothesized for skill-based constructs. However, evidence of the written communication higher-order factor structure was supported. Substantial focal path coefficients ( $\beta > 0.500$ ) in which the higher-order factor (written communication) significantly predicted the lower-order factors (supervisor, peer, subordinate, and other). The written communication higher-order factor accounted for 25 percent of the variance in the model, which is higher than the other six competencies.

Table 36

## Structural Model Results: Written Communication

Relationships	Hypothesized Model
LT → WRIT	-0.289
SR → WRIT	n.s.
SS → WRIT	n.s.
FF → WRIT	0.371
EE → WRIT	-0.366
WRIT → SUPERVISOR	0.569
WRIT → PEER	0.675
WRIT → SUBORDINATE	0.462
WRIT → OTHER	0.504
Model Fit Indices	
$\chi^2$	1101.6
<i>df</i>	721
CFI	0.885
RMSEA	0.055
Squared Multiple Correlation (WRIT HOF)	0.245

*Note.* All chi-squares were significant at  $p < 0.001$ . n.s. = not significant, EE = Emotional Evenness, FF = Need to be like, LT = Liking for Thinking, SR = Self-Reliance, SS = Sociability, WRIT = Written Communication (HOF = Higher-Order Factor).

## **CHAPTER FOUR**

### **DISCUSSION**

This research provided empirical evidence demonstrating the partial construct validity for combining 360-degree feedback competency ratings into single construct measures of performance using confirmatory factor analysis. Mixed results were found for the broad and specific hypothesized personality relationships of the Big Five to trait-based and skill-based 360-degree feedback competency constructs using structural equation modeling. Previous research has primarily focused on how raters in 360-degree feedback systems tend to systematically agree or disagree within or between-groups (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988). However, this research extended Schmidt and Hunter's (1996) proposition that inter-rater agreement can be interpreted as construct reliability coefficients by treating raters as passive instruments. Although interpreting 360-degree feedback competency ratings in this manner may not make sense for developmental feedback, the results of this study support the notion that by combining ratings into constructs, construct validity evidence is obtainable and 360-degree feedback competency ratings can be used as a criteria measure of performance. This study provides additional evidence regarding the debate over whether to use multiple criteria versus composite criteria when measuring 360-degree feedback competency ratings.

As reported in Figure 1, 360-degree feedback ratings were measured from multiple perspectives, including the self, supervisor, peers, subordinates, and others. London and Smither (1995) provided evidence that self-ratings tend to be poor indicators of true performance and have the highest level of disagreement with other types of raters. The present research included models containing self-ratings and excluding self-ratings in an effort to give perspective into the role self-rating error may play in 360-degree composite criteria. As a first step, separate CFA models were created to examine the seven 360-degree feedback competencies with averaged ratings across all rater types. This created a baseline of how to understand and interpret competency construct validity and fit. When the differences between the observed model and measurement model were compared, statistically significant differences were found in the chi-squares of the seven models, indicating the measurement models were different than the observed models. Also, the fit indices indicated that the CFA models did not meet the goodness-of-fit criteria set out by Bagozzi and Yi (2012), but the models did fall within Bagozzi and Yi's (2012) badness-of-fit criteria. Simply stated, the models did not fit well, but they also did not fit poorly.

Consistent with previous research on the unreliability of self-ratings (Jones et al., 1981; London & Smither, 1995; Thornton, 1980), the current study found that self-ratings consistently had lower factor loadings across all seven competencies than did other rater types (i.e., peer ratings, supervisor ratings, etc.). Also, self-ratings did not meet the established .50 factor loading criteria set by Bagozzi and Yi (2012). However, creating CFA models with ratings averaged across rater type and including self-ratings was not found to be the optimal way of constructing ratings. The models lacked convergent

validity evidence outlined by Hair et al. (2006) of having over 50 percent variance extracted and construct reliability ratings over .70. All seven of the constructs lacked convergent validity evidence, which led the researcher to pursue other hypothesized models. Although it was not directly related to the proposed hypotheses of this research, it was noted that of the seven competencies measured, none were significantly superior or inferior to the others in terms of fit or validity. This finding is important because it supports the notion that, psychometrically, competencies can be used as similar measures of performance regardless of what the competency intended to measure. Consequently, when using competencies in future research, results should be similar between confirmatory or structural models, regardless of what competency is measured.

In accordance with the second hypothesis, the researcher altered the previous seven CFA models by simply excluding the self-ratings. Observed and measurement model differences were compared to Hypothesis One, along with the fit indices necessary for all seven CFA models. When self-ratings were excluded from the CFA models, alignment between the measurement and observed models increased (i.e., decreased chi-squares) and fit indices increased across all seven models, which provided support for Hypothesis Two. This replicated and confirmed previous research findings (London & Smither, 1995; Warr & Hoare, 2002) that self-ratings tend to be error-laden measures of performance within 360-degree feedback systems. By excluding self-ratings from future confirmatory and structural models of 360-degree feedback competencies, higher reliability and validity coefficients are more likely. Also, by excluding self-ratings from 360-degree feedback constructs, these constructs could serve as composite criteria in future predictive models, and dependent variables for future predictors, such as selection

predictors, turnover, and high-potential identification. Also, previous theories of job performance could be modeled by substituting 360-degree feedback constructs for performance measures. However, there were still flaws found in the Hypothesis Two approach to constructing 360-degree feedback ratings. By averaging across rater-types, it was proposed that much of the valuable rater variance would be lost, even when error-prone self-ratings were excluded. The loss in variance among the raters may have suppressed the observed convergent validity evidence in the Hypothesis Two analyses. Consequently, other hypotheses were explored for models that had better fit indices while avoiding the loss in rater variance.

Hypothesis Three stated that the five personality scales of the GZTS/DFOS personality inventory, which corresponded to the five scales of the Big Five, would present acceptable fit indices and construct validity evidence. Although some previous research indicated that a lack of dimensionality can be found when running confirmatory models of the lexical, Big Five theory of personality (Hopwood & Donnellan, 2010), the current study provides evidence that the five scales of the GZTS/DFOS personality inventory, which corresponded to the Big Five, presented acceptable fit and mixed, but inconclusive evidence of construct validity. As was stated in the results, Hair et al. (2006) stated for construct validity to be supported, convergent, discriminant, nomological, and face validity evidence must all be found. For the personality scales in this research, only convergent and discriminant validity were possible to examine, while information on nomological and face validity were not available due to the confidentiality agreements necessary to obtain the data. Nomological validity (i.e., whether item correlations with other items or item factor loadings make logical sense) and face validity (i.e., whether an

item's content judgmentally excludes the item from being classified under a particular construct) were not investigated because the researcher did not have access to item content. Convergent validity was supported for one of the five constructs, while discriminant validity was supported for all five of the constructs. Partial evidence of construct validity was found for the GZTS/DFOS personality inventory scales related to the Big Five. Because of the positive, yet inclusive, convergent and discriminant validity findings, the researcher was able to move forward and test a comprehensive, hierarchical model of 360-degree feedback.

Problems with acceptable fit indices were found in the previous hypotheses that may have been due to the averaging procedures used across rater types. Hypothesis Four proposed a hierarchical model of 360-degree feedback competencies (seen in Figure 6), where competencies served as higher-order constructs and rater types served as lower-order constructs (while individual ratings served as items). By structuring 360-degree feedback competencies this way, averaging procedures were not necessary and more of the variance was accounted for across all rater types, items, and factor loadings. However, when Schmidt and Hunter (1996) proposed arranging 360-degree feedback ratings to serve as intercorrelations between ratings as a measure of construct validity, they did not propose arranging 360-degree feedback competencies hierarchically. Consequently, the present research based the theoretical necessity of constructing ratings hierarchically on the amount of rater variance that would be accounted for by this type of model.

Hypothesis Four stated that constructing the 360-degree feedback competencies hierarchically would increase the fit indices and present more construct validity evidence

than by averaging ratings across rater types. When looking at differences between the observed and measurement models, large increases (chi-squares doubling or quadrupling) were found in the differences between the observed and measurement models and the number of degrees of freedom (i.e., the amount of mathematical information available to estimate model parameters). Moreover, the increased differences were most likely due to the substantial complexity found in the models, along with the increase of sample moments available from which to draw degrees of freedom. Although the findings of increased model differences could be negative, the increased number of degrees of freedom adds power to the model. However, the fit indices related to hierarchical structure saw improvements in fit, with some competencies meeting Hair et al.'s (2006) criteria for CFI and RMSEA fit. Because of the improvements in fit, and the increase in the number of degrees of freedom by using the hierarchical models, the determination was made that hierarchical models of 360-degree feedback competencies were superior to models averaging across rater types from Hypotheses One and Two.

Evidence of convergent and discriminant validity was found for constructing the models hierarchically as well. However, for Hypothesis Four to be supported as the best construction of 360-degree feedback competency ratings, fit indices and validity evidence must be compared. The comparison of the model fit and convergent validity evidence between Hypothesis One, Two, and Four's different configurations of 360-degree feedback, are reported in Table 29. From this table, the researcher concluded that the CFA models that met the most criteria of construct validity were the hierarchical models of Hypothesis Four. Consequently, all of the hierarchical CFA models were transformed into structural models outlined by Hypothesis Five and Six.

One of the central propositions of this research was testing the differences between trait-based and skill-based 360-degree feedback competencies. The present research hypothesized that personality variables of the Big Five would significantly predict global relationships with trait-based competencies as well as some specific personality-competency relationships, but the personality variables would not significantly predict any relationships with skill-based competencies.

The findings for Hypothesis Five (personality predicting trait-based competencies), were mixed in that no significant relationships were found between personality and trait-based constructs for two (i.e., Drive for Results and planning and organizing) of the five competencies classified as *trait-based*. Also, some of the personality-competency relationships were found to be significant but negative in direction. Specifically, the surrogate personality measure for Openness to Experience was found to be a negative predictor of both relationship management and managing others. This finding suggests that people who are open to new experiences would be rated by others as a poor performer in terms of how they manage others and how they manage their relationships. Perhaps, managers who rate themselves highly in Openness to Experience focus more on new opportunities, rather than directing and leading others on present opportunities and/or developing and maintaining positive work relationships with others.

Other specific findings for personality as a predictor of trait-based competencies included the surrogate for Extroversion predicting relationship management and the surrogate for Emotional Stability predicting resilience. Both of these findings supported the specific hypothesized relationships between personality and 360-degree feedback

competencies. However, it was noted that of the eleven specific personality-competency relationships hypothesized, only these two specific personality-competency relationships were statistically significant. The limited number of specific relationships is concerning because of some of the semantic similarities between the personality traits and the competencies. When creating the specific hypothesized relationships, it was difficult for the researcher to determine which relationships would be positive because there was very little information available describing each competency and personality trait.

Models which measured the null relationships proposed in Hypothesis Six between the Big Five personality traits and skill-based 360-degree feedback competencies were not supported. Null personality-competency relationships were expected across the two competencies being measured: business acumen and written communication. However, the surrogate for conscientiousness was found to significantly predict business acumen, and the surrogates for Openness to Experience and Emotional Stability were found to negatively predict written communication, while Agreeableness was found to positively predict written communication. Moreover, the three predictors of written communication had the most significant predictive relationships of all the personality-competency relationships found in any of the seven hierarchical models hypothesized.

These findings were baffling and beg the question as to why the relationship between personality and skill-based competencies was significant. These findings could possibly be statistical artifacts or errors of measurement. However, when being rated by others, an important aspect of being rated highly for having skills in business acumen could be interpreted as having a high predisposition for being Conscientious.

Overall, Hypothesis Five received mixed support, while Hypothesis Six was not supported at all. As mentioned earlier, the lack of information provided to the researcher about the 360-degree feedback item content could have played a role in the improper classification of 360-degree feedback competencies as either being trait- or skill-based. Although at face value, Drive for Results may be seen as a construct related to one's individual differences (e.g. drive, commitment, etc.), it is possible that when being rated by others, it could be seen as more related to one's skills or behaviors necessary for success. Similarly, planning and organizing could be seen as a construct that is most likely related to personality characteristics, such as conscientiousness. However, it is possible that when others are rating planning and organizing, the relationships could be seen as being less-related to conscientious traits and more-related to the skills associated with scheduling or performance. Conversely, although business acumen could be seen as the skills necessary to understand general business and financial concepts, certain personality characteristics could be responsible for those skills, such as conscientiousness. Moreover, although written communication could come from years of writing experience and previous writing skills, it may also be negatively related to Openness to Experience and Emotional Stability, while being positively related to Agreeableness.

Perhaps, the semantic classification system by which competencies were categorized by the researcher as being more trait-based or more skill-based was flawed. Semantically, business acumen and written communication may have consisted of components that were more associated with trait-based competencies, while Drive for Results and planning and organizing, although seemingly related to personality

characteristics, may have been more skill-based in the context of 360-degree feedback. Notwithstanding, the previously classified trait-based competencies, Drive for Results and planning and organizing, may have still possibly been related to personality or individual difference characteristics, but not those characteristics measured by the Big Five. Another alternative explanation could be that classifying 360-degree feedback competencies into Spencer and Spencer's (1993) trait-based and skill-based philosophical perspectives was not a theoretically sound model of constructing hierarchical structural models of 360-degree feedback.

DeNisi and Kluger's (2000) findings that feedback directed toward individual difference characteristics of the *ideal self* (i.e., inborn predispositions or trait-based competencies) are related to negative outcomes, such as low self-esteem, self-efficacy, and productivity, could be a concern of this research. The current study found significant relationships between some stable, individual difference variables of personality and *changeable* 360-degree feedback competencies. These findings may concern Self-Regulatory theorists, such as DeNisi and Kluger, because if these significant relationships hold true, 360-degree feedback programs related to personality may not be as effective as expected. The finding that feedback competencies may be relatively stable and may be difficult to develop could be why Smither et al. (2005) stated that organizations "should not expect large, widespread performance improvements after employees receive multi-source feedback" (p. 33). However, more research into the effects of personality relationships to 360-degree feedback competencies is needed to investigate the personality-competency relationship with managerial outcomes.

Methodologically, however, the present research provided evidence that using hierarchical models of 360-degree feedback competencies can illuminate the shortcomings of the averaging techniques currently used when analyzing 360-degree feedback ratings. Whether measuring trait-based or skill-based competencies, all seven of the higher-order factor constructs were significantly predictive of lower-order factors with noticeably high regression coefficients (above .500). However, other than written communication (with 25 percent variance accounted for), all of the higher-order factor competencies accounted for less than nine percent of the variance. These findings, alongside the fit indices and construct validity evidence, encourage future debates over constructing 360-degree feedback competencies in hierarchical CFA and SEM models. However, further evidence is needed for assembling 360-degree feedback competencies hierarchically. Future research should focus on building construct valid hierarchical models, along with assessing the nomological and face validity evidence. Also, future research should focus on assessing alternative models to those presented within this research.

### **Limitations**

Sample sizes within each confirmatory and structural model ranged from  $n = 177$  to having over a thousand participants per construct, indicating great sample size differences among the seven competencies. By having such a range in sample sizes, claims of generalizability between competencies made the comparisons strained. With some competencies having low sample sizes, attempts to assess construct validity were more restricted.

Another limitation of this research was the researcher's lack of access to the item content of the personality scales and 360-degree feedback competencies. Not having this information made hypothesizing specific and broad personality-competency relationships based on previous theoretical research impossible. As a result, some spurious findings (e.g., personality traits predicting 360-degree feedback competencies which were not semantically related) were anticipated in exploratory research of this nature.

When conducting exploratory research, such as constructing hierarchical structural models of 360-degree feedback, the lack of prior research and unclear theoretical implications are limitations. No previous research examples of constructing 360-degree feedback in CFA or SEM models were found. Thus, the prior theoretical and methodological foundation for constructing 360-degree feedback in the manner presented in this research was not available.

Lastly, by using a cross-sectional convenience sample, certain limitations may have been present, such as common method variance and having a common sample pool.

### **Suggestions for Future Research**

Additional research into the methodological effects of using 360-degree feedback competency ratings in both confirmatory and structural models is needed. Also, it is recommended that future research include higher sample sizes and equality between samples. This would help to improve comparisons made between competencies. Future research could seek to replicate this study with broader samples of 360-degree feedback interventions within organizations (e.g. executive development or differing industries) other than just being used for middle-management development. Data could also be collected longitudinally to compare 360-degree feedback results and relationships as they

develop over time. Factors such as culture, demographics, and other mediators could all play a role in 360-degree feedback in future research.

Creating a superordinate model of 360-degree performance ratings with predictors, mediators, and criteria all theorized into a hierarchical structural model could advance the field of personnel research considerably. By creating such a model, constructs, such as leadership, citizenship behaviors, cognitive ability, and others, could be modeled as a comprehensive framework of performance for selecting and promoting managers or executives. Perhaps, by organizing 360-degree feedback ratings hierarchically, 360-degree feedback performance ratings could serve as an alternative and valid measure of managerial performance in the future.

## REFERENCES

- Abdullah, A. H., Musa, R. M., & Ali, J. H. (2011). The development of human resource practitioner competency model perceived by Malaysian human resource practitioners and consultants: A structural equation modeling approach. *International Journal of Business and Management, 6*(11), 240-257.
- Aguinis, H. (2008). Enhancing the relevance of organizational behavior by embracing performance management research. *Journal of Organizational Behavior, 29*, 315-342.
- Ammons, R. B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. *Journal of General Psychology, 54*, 279-299.
- Antonioni, D. (1996). Designing an effective 360-degree appraisal feedback process. *Organizational Dynamics, 6*, 157-171.
- Ashford, S. J., & Cummings, L. L. (1983). Feedback as an individual resource: Personal strategies of creating information. *Organizational Behavior and Human Performance, 32*, 370-398.
- Atwater, L. E., & Brett, J. (2006). Feedback format: Does it influence manager's reactions to feedback? *Journal of Occupational and Organizational Psychology, 79*, 517-532.
- Atwater, L. E., Rouch, P., & Fischthal, A. (1995). The influence of upward feedback on self and follower raters of leadership. *Personnel Psychology, 48*, 34-60.

- Atwater, L. E., & Waldman, D. (1998). 360-degree feedback and leadership development. *Leadership Quarterly*, 9, 423-426.
- Atwater, L. E., Waldman, D., & Brett, J. (2002). Understanding and optimizing multi-source feedback. *Human Resource Management Journal*, 41, 193-208.
- Avery, K. (2000). The effects of 360-degree feedback over time. (Unpublished doctoral dissertation). Chicago School of Professional Psychology.
- Bagozzi, R. P. (1994). Measurement in marketing research. In R. P. Bagozzi (Eds.), *Principles in Marketing Research*, (pp. 1-49). Cambridge, MA: Blackwell.
- Bagozzi, R. P., & Yi, Y.(2012).Specifications, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, 40, 8-34.
- Bailey, C., & Austin, M. (2006). 360-degree feedback and developmental outcomes: The role of feedback characteristics, self-efficacy and importance of feedback dimensions to focal managers' current role. *International Journal of Selection and Assessment*, 14, 51-66.
- Bailey, C., & Fletcher, C. (2002) *When do other peoples' opinions matter? The credibility of feedback from co-workers*. Presented at the annual meeting of the British Psychological Society, Blackpool, UK.
- Banasova, L., Caganova, D., & Cambal, M. (2011). The identification of key managerial competencies as a tool for increasing business competitiveness. *European Conference of Knowledge Management*, 2, 46-53.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.

- Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.
- Bartram, D. (2005). The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 90*, 1185-1203.
- Becker, B. E., & Huselid, M. A. (1999). Overview: Strategic human resource management in five leading firms. *Human Resource Management, 38*, 287-301.
- Becker, T. E., & Klimoski, R. J. (1989). A field study of the relationship between organizational feedback environment and performance. *Personnel Psychology, 42*, 343-358.
- Beheshtifar, M., & Moghadam, M. N. (2011). Studying the competency-related models in succession planning. *European Journal of Economics, Finance, and Administrative Sciences, 34*, 113-121.
- Bem, D. J. (1972). Self-Perception Theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 6, pp. 1-62). New York: Academic Press.
- Bentler, P. M. (1985). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.
- Bentler, P. M., & Bonnett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606.
- Bernardin, H. J., & Beatty, R. W. (1984). *Performance Appraisal: Assessing Human Behavior at Work*. Boston, MA: Kent Publishing.

- Bernardin, H. J., & Beatty, R. W. (1987). Can subordinate appraisals enhance managerial productivity? *Sloan Management Review*, 28(4), 63-73.
- Bernardin, H. J., Dahmus, S. A., & Redmon, G. (1993). Attitudes of first-line supervisors toward subordinate appraisals. *Human Resource Management*, 32, 315-324.
- Bettenhausen, K., & Fedor, D. (1997). Peer and upward appraisals: A comparison of their benefits and problems. *Group and Organization Management*, 22, 236-263.
- Bliszczyk, R., & Dimasi, V. (2003). 360-degree feedback: An Australian study. *Australian Journal of Psychology*, 55, 115-116.
- Blum, M. L., & Naylor, J. C. (1968). *Industrial Psychology: Its Theoretical and Social Foundations* (3<sup>rd</sup> Edition). New York: Harper & Row.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.
- Borman, W. C. (1974). The rating of individuals in organizations: An alternative approach. *Organizational Behavior and Human Performance*, 12, 105-124.
- Borman, W. C. (1997). 360-degree ratings: An analysis of assumptions and a research agenda for evaluating their validity. *Human Resource Management Review*, 7, 299-315.
- Bozeman, D. P. (1997). Inter-rater agreement in multi-source performance appraisal: A commentary. *Journal of Organizational Behavior*, 18, 313-316.
- Bracken, D. W. (1994). Straight talk about multi rater feedback. *Training and Development*, 48, 44-50.
- Bracken, D. W. (2006). Maximizing the uses of multi-rater feedback. *Center for Creative Leadership*, 12, 11-17.

- Bracken, D., & Timmreck, C. (2001). Guidelines for multi-source feedback when used for decision-making. In D. Bracken & C. Timmreck (Eds.), *The Handbook of Multi-source Feedback*. San Francisco: Jossey-Bass.
- Bracken, D. W., Timmreck, C. W., & Church, A. H. (2001). *The Handbook of Multi-source Feedback: The Comprehensive Resource for Designing and Implementing MSF Processes*. San Francisco, CA: Jossey-Bass.
- Brannick, M. T., Levine, E. L., & Morgeson, F. P. (2007). *Management and Teams. Job and Work Analysis: Methods, Research, and Applications for Human Resource Management* (2<sup>nd</sup> Edition). Los Angeles, CA: Sage.
- Brett, J., & Atwater, L. (2001). 360-degree feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Management Development*, 18(8), 676-693.
- Brewster, C., Farndale, E., & Ommeren, J. V. (2000). HR competencies and professional standards. *World Federation of Personnel Management Associations*, CIPD, UK.
- Brusman, M. (2008). Multi-rater 360-degree feedback. *Working Resources*, 25, 115-132.
- Brutus, S., Fleenor, J. W., & McCauley, C. D. (1999). Demographic and personality predictors of congruence in multi-source ratings. *Journal of Management Development*, 18, 417-435.
- Bryk, A. S., & Raudenbush, S. W. (2001). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage.
- Buda, R., Reilly, R. R., & Smither, J. W. (1991). The influence of indirect knowledge of prior performance on evaluations of present performance: The generalizability of assimilation effects. *Journal of Psychology and the Behavioral Sciences*, 6, 89-99.

- Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology, 90*, 60-77.
- Caputo, P. M., & Roch, S. (2009). *Rating formats and perceptions of performance appraisal fairness*. Poster presented at the annual meeting of the Society for Industrial and Organizational Psychology, New Orleans, LA.
- Cardy, R. L., & Dobbins, G. H. (1994). *Performance Appraisal: Alternative Perspectives*. Cincinnati, OH: South-Western Publishing.
- Carl, D., Gupta, V., & Javidan, M. (2004). Power distance. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta (Eds.), *Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies* (pp. 433-446). San Francisco: Jossey-Bass.
- Carruthers, F. (2003). Nothing but the truth. *Australian Financial Review, 14*, 78-88.
- Carver, C. S. (2007). Approach, avoidance, and the self-regulation of affect and action. *Motivation and Emotion, 30*, 105-110.
- Carver, C. S., & Scheier, M. F. (1981). *Attention and Self-regulation: A Control Theory Approach to Human Behavior*. New York: Springer-Verlag.
- Chappelow, C. T. (2004). 360-degree feedback. In C. D. McCauley & E. Van Velsor (Eds.), *The Center for Creative Leadership Handbook of Leadership Development* (2nd ed., pp. 58-84). San Francisco: Jossey-Bass.
- Cheung, G. W. (1999). Multifaceted conceptions of self-other ratings disagreement. *Personnel Psychology, 52*, 1-36.

- Church, A. H. (1997). Do you see what I see? An exploration of congruence in ratings from multiple perspectives. *Journal of Applied Social Psychology, 27*, 983-1020.
- Coens, T., & Jenkins, M. (2000). *Abolishing Performance Appraisals: Why They Backfire and What to do Instead*. Berrett-Koehler Publishers.
- Conway, J. M. (1996). Analysis and design of multi-trait multi-rater performance appraisal studies. *Journal of Management, 22*, 139-162.
- Conway, J. M., & Huffcutt, A. E. (1997). Psychometric properties of multi-source feedback ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance, 10*, 331-360.
- Costa, P. T., & McCrae, R. R. (1992). *NEO Personality Inventory-Revised (NEO-PI-R) and the NEO Five-Factor Inventory (NEO-FFI) Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.
- Costa, P. T., Jr., & McCrae, R. R. (1997). Stability and change in personality assessment: The revised NEO personality inventory in the year 2000. *Journal of Personality Assessment, 68*, 86-94.
- Cronshaw, S. F., & Lord, R. G. (1987). Effects of categorization, attribution, and encoding processes of leadership perceptions. *Journal of Applied Psychology, 72*, 97-106.
- Dai, G., De Meuse, K. P., & Peterson, C. (2010). Impact of multi-source feedback on leadership competency development: A longitudinal field study. *Journal of Management Issues, 22*, 197-219.
- Dalton, M. (1997). When the purpose of using multi-rater feedback is behavior change. D. Bracken, M. Dalton, R. Jako, C. McCarthy, and V. Pollman, (Eds.), *Should*

*360-Degree Feedback Be Used Only for Development Purposes?* Greensboro, NC: Center for Creative Leadership.

DeNisi, A. S., & Kluger, A. N. (2000). Feedback effectiveness: Can 360-degree appraisals be improved? *Academy of Management Executive*, *14*, 129-139.

Dominick, P. G., Reilly, R. R., & Byrne, J. C. (2004). *Individual differences and peer feedback: Personality's impact on behavior change*. Paper presented at the annual conference of the Society for Industrial and Organizational Psychology. Chicago, Illinois.

Drucker, P. F. (1973). *Management: Tasks, Responsibilities, Practices* (Drucker Series). Los Angeles, CA: Harper-Business.

Eckert, R., Ekelund, B. Z., Gentry, W. A., & Dawson, J. F. (2010). "I don't see me like you see me, but is that a problem?" Cultural influences on rating discrepancy in 360-degree feedback instruments. *European Journal of Work and Organizational Psychology*, *19*, 259-278.

English, A., Rose, D., & McClellan, J. (2009). *Rating scale label effects on leniency bias in 360-degree feedback*. Paper presented at the annual conference of the Society of Industrial and Organizational Psychology, New Orleans, LA.

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

Farh, J. L., Werbel, J. D., & Bedeian, A. G. (1988). An empirical investigation of self-appraisal-based performance evaluation. *Personnel Psychology*, *41*, 141-156.

- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Fitzgibbons, A. (2003). *Examining the equivalence of rater groups in 360-degree feedback for use in leadership development*. Retrieved from ProQuest Dissertations and Theses. (3126511).
- Fletcher, C. (1999). The implication of research on gender differences in self-assessment and 360-degree appraisal. *Human Resource Management Journal*, 9, 39-46.
- Fletcher, C., & Baldry, C. (2000). A study of individual differences and self-awareness in the context of multi-source feedback. *Journal of Occupational and Organizational Psychology*, 73, 303-319.
- Funderburg, S. A., & Levy, P. E. (1997). The influence of individual and contextual variables on 360-degree feedback system attitudes. *Group and Organizational Management*, 22, 210-235.
- Glasgow, L. P. (1999). A comparison of broad and narrow personality traits in the prediction of job performance. (Unpublished doctoral dissertation). Southern Methodist University, Dallas, TX.
- Goldsmith, M., & Underhill, B. (2001). Multi-source feedback for executive development. In D. Bracken, C. Timmreck, & A. Church (Eds.), *The Handbook of Multi-source Feedback* (pp. 275-288). San Francisco: Jossey-Bass.
- Gray, J. A. (1994). Personality dimensions and emotion systems. *The Nature of Emotion: Fundamental Questions* (Ed.), (pp. 329-331). San Francisco: Jossey-Bass.
- Greguras, G. J., & Robie, C. (1998). A new look at within-source inter-rater reliability of 360-degree feedback ratings. *Journal of Applied Psychology*, 83, 960-968.

- Guilford, J. P., Christensen, P. R., & Bond, N. A. (1956). *The DF Opinion Survey: Manual of Instructions and Interpretations*. Beverley Hills, CA: Edits Publishing.
- Guilford, J. P., & Martin, H. G. (1944). Age differences in some introvertive and emotional traits. *Journal of General Psychology*, *31*, 219-229.
- Guilford, J. P., & Zimmerman, W. S. (1949). *The Guilford-Zimmerman Temperament Survey: Manual*. Beverly Hills, CA: Sheridan Supply Co.
- Guilford, J. P., Zimmerman, W. S., & Guilford, J. S. (1976). *The Guilford-Zimmerman Temperament Survey Handbook*. San Diego: Edits Publishing.
- Hackman, J. R., & Oldham, G. R. (1980). *Work Redesign*. Reading, MA: Addison-Wesley.
- Hagan, C. M., Konopaske, R., Bernardin, H. J., & Tyler, C. L. (2006). Predicting assessment center performance with 360-degree, top-down, and customer-based competency assessments. *Human Resource Management*, *45*(3), 357-390.
- Hair Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate Data Analysis*. (6<sup>th</sup> Edition). Pearson Prentice Hall. Upper Saddle River, NJ.
- Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, *41*, 43-62.
- Hassan, S., & Rohbaugh, J. (2009). Incongruity in 360-degree feedback ratings and competing managerial values: Evidence from a public agency setting. *International Public Management Journal*, *12*, 421-449.
- Hernez-Broome, G., & Hughes, R. L. (2004). Leadership development: Past, present, future. *Human Resource Planning*, *27*, 24-32.

- Hezlett, S. A., & Ronkvist, A. M. (1996). *The effects of multi-rater feedback on managers' skill development: Factors influencing behavior change*. Paper presented at the annual conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319-340.
- Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values*. Thousand Oaks, CA: Sage.
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations*. Thousand Oaks, CA: Sage.
- Hogan, R. (2004). Personality psychology for organizational research. *Personality and Organizations*, *4*, 3-24.
- Hopwood, C.J., & Donnellan, M.B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, *14*, 332-346.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (3rd Ed., pp. 114-157). New York, NY: Guilford Press.
- Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*, 1-2.

- Jones, E. E., Rhodewalt, F., Berglas, S., & Skelton, J. A. (1981). Effects of strategic self-presentation on subsequent self-esteem. *Journal of Personality and Social Psychology, 41*, 407-421.
- Judge, T. A., Bono, J. E., Ilies, R., & Gebhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*, 765-780.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-291.
- Kanfer, F. H. (1980). Self-management methods. *Helping People Change: Methods* (2nd Ed., pp. 334-389). Thousand Oaks, CA: Sage.
- Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology, 44*, 23-52.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: Historical review, a meta-analysis, and a preliminary Feedback Intervention Theory. *Psychological Bulletin, 119*, 254-284.
- Korman, A. K. (1970). Toward a hypothesis of work behavior. *Journal of Applied Psychology, 54*, 31-41.
- Krugar, J., & Dunning, D. (2002). Unskilled and unaware—but why? A reply to Krugar and Mueller. *Journal of Personality and Social Psychology, 82*, 189-192.
- Latham, G. P., & Wexley, K. N. (1994). *Increasing Productivity through Performance Appraisal* (2nd Ed.). Reading, MA: Addison-Wesley.
- Lawler, E. E., III. (1967). The multitrait-multirater approach to measuring managerial job performance. *Journal of Applied Psychology, 51*, 369-381.

- Leslie, J. B. (2002). *360-degree Feedback: Best Practices to Ensure Impact*. Greensboro, NC: Center for Creative Leadership.
- Levy, P. E., Cawley, B. D., & Foti, R. J. (1998). Reactions to appraisal discrepancies: Performance ratings and attributions. *Journal of Business and Psychology, 12*, 437-455.
- Lin, Y. (2012). *Predicting 360 ratings with personality: Comparing linear and curvilinear models*. Paper presented at the annual conference of the Society of Industrial and Organizational Psychology, San Diego, CA.
- Locke, E. A. (1996). Motivation through conscious goal setting. *Applied and Preventative Psychology, 5*, 117-124.
- London, M., Wohlers, A. J., & Gallagher, P. (1990). 360-degree feedback surveys: A source of feedback to guide management development. *Journal of Management Development, 9*, 17-31.
- London, M., & Beatty, R. W. (1993). 360-degree feedback as a competitive advantage. *Human Resource Management, 32*, 353-372.
- London, M., & Smither, J. W. (1995). Can multi source feedback change perception of goal accomplishment, self-evaluations, and performance-related outcomes? Theoretical applications and directions for research. *Personnel Psychology, 48*, 803-840.
- Longenecker, C., & Ludwig, D. (1990). Ethical dilemmas in performance appraisals revisited. *Journal of Business Ethics, 9*, 961-969.

- Lustri, D., Miura, I., & Takahashi, S. (2007). Knowledge management model: Practical application for competency development. *The Learning Organization, 14*, 186-201.
- Mabe, P. M., & West, S. G. (1982). Validity of self-evaluations of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*, 280-296.
- Markus, H., & Sentis, K. (1982). The self in social information processing. In J. Suls (Ed.), *Social Psychological Perspectives on the Self* (pp. 154-167). Hillsdale, NJ: Erlbaum.
- Maurer, T. J., Barbeite, F. G., & Mitchell, D. R. (2002). Predictors of attitudes towards a 360-degree feedback system and involvement in post-feedback management development activity. *Journal of Occupational & Organizational Psychology, 75*, 87-107.
- Maurer, T. J., Raju, N. S., & Collins, W. C. (1998). Peer and subordinate performance appraisal measurement equivalence. *Journal of Applied Psychology, 83*, 693-702.
- Maylett, T., & Riboldi, J. (2007). Using 360-degree feedback to predict performance. *Training & Development, 61*(9), 87-101.
- McCall, M. W., & Lombardo, M. M. (1983). *Off the Track: Why and How Successful Executives get Derailed*. Greensboro, NC: Center for Creative Leadership.
- McCarthy, A. M., & Garavan, T. N. (2001). 360-degree feedback processes: Performance improvement and employee career development. *Journal of European Industrial Training, 25*, 3-32.
- McCrae R. R., Costa P. T., Martin T. A. (2005). The NEO-PI-3: A more readable revised NEO personality inventory. *Journal of Personality Assessment, 84*, 261–270.

- McEvoy, G. M., & Buller, P. F. (1987). User acceptance of peer appraisals in an industrial setting. *Personnel Psychology, 40*, 785-797.
- McPherson, B. (2007). Does your staff think you're a better manager? Evaluating leadership development. *The International Journal of Leadership in Public Services, 3*(4), 69-72.
- Mersman, J., & Donaldson, S. (2000). Factors affecting the convergence of self-peer rating in contextual and task performance. *Human Performance, 13*, 299-322.
- Meyer, H. H. (1991). A solution to the performance appraisal feedback enigma. *Academy of Management Executive, 5*, 68-76.
- Miller, J. S., & Cardy, R. L. (2002). Self-monitoring and performance appraisal: Rating outcomes in project teams. *Journal of Organizational Behavior, 21*, 609-626.
- Mitchell, T. (1979). Organizational Behavior. In M. R. Rosenzweig & I. W. Porter (Eds.) *Annual Review of Psychology* (pp. 243-281). Palo Alto, CA: Annual Reviews.
- Mitchell, T. R., & Daniels, D. (2003). Motivation. Comprehensive handbook of psychology: *Industrial Organizational Psychology, 12*, 225-254.
- Mittal, R., & Saran, A. (2010). Impact of power distance on multi-source feedback evaluations: A cross-cultural perspective. *Journal of International Business and Economics, 10*, 117-122.
- Morgeson, F. P., Mumford, T. V., & Campion, M. A. (2005). Coming full circle: Using research to address 27 questions about 360-degree performance ratings. *Consulting Psychology Journal: Practice and Research, 57*, 196-209.

- Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. *Research in Personnel and Human Resource Management, 13*, 153-200.
- Mount, M. K., Judge, T. A., Scullen, S. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. *Personnel Psychology, 51*, 557-569.
- Murphy, K. R., & Cleveland, J. N. (1991). *Performance Appraisal: An Organizational Perspective*. Boston, MA: Allyn & Bacon.
- Murphy, K., & Cleveland, J. (1995). *Understanding Performance Appraisal: Social, Organizational, and Goal-Oriented Perspectives*. Thousand Oaks, CA: Sage.
- Nadler, D. A. (1977). *Feedback and Organizational Development: Using Data Based Methods*. Reading, MA: Addison-Wesley.
- Nemeroff, W. F., & Cosentino, J. (1979). Utilizing feedback and goal setting to increase performance appraisal interview skills of managers. *Personnel Psychology, 51*, 557-569.
- Nilsen, D., & Campbell, D. (1993). Self-observer rating discrepancies: Once an over-rater, always an over-rater? *Human Resource Management, 32*, 265-281.
- Nowack, K. (1992). Self-assessment and rater assessment as a dimension of management development. *Human Resources Quarterly, 3*, 141-155.
- Nowack, K., & Mashihi, S. (2012). *Evidence based answers to ten questions about leveraging 360-degree feedback*. Paper presented at the annual conference of the Society for Industrial and Organizational Psychology, San Diego, CA.

- Nutt, P. (1993). The formulation processes and tactics used in organizational decision-making. *Organizational Science, 4*, 226-251.
- Ostroff, C., Atwater, L. E., & Feinburg, B. J. (2004). Understanding self-other agreement: A look at rater characteristics, context, and outcomes. *Personnel Psychology, 57*, 333-375.
- Paterson, D. G. (1922). The Scott Company graphic rating scale. *Journal of Personnel Research, 1*, 361-376.
- Pollitt, D. (2005). Leadership succession planning affects commercial success. *Human Resource Management International Digest, 13*, 36-38.
- Pollman, V. (1997). Some faculty assumptions that support using multi-rater feedback for performance appraisal. In D. Bracken, M. Dalton, R. Jako, C. McCauley, & V. Pollman (Eds.) *Should 360-degree Feedback be used Only for Developmental Purposes?* (pp.7-9). Greensboro, NC: Center for Creative Leadership.
- Rabinowitz, I., Kelley, H., & Rosenblatt, R. (1966). Effects of different types of interdependence and response conditions in the minimal social situation. *Journal of Experimental Social Psychology, 2*, 169-197.
- Rainy, H. G. (1997). *Understanding and Managing Public Organizations*. San Francisco: Jossey-Bass.
- Reilly, R. R., & McGourty, J. (1998). Performance appraisal in team settings. In J.W. Smither (Ed.), *Performance Appraisal: State of the Art in Practice* (pp. 53-61). San Francisco: Jossey-Bass.
- Reilly, R. R., Smither, J. W., & Vasilopoulos, N. L. (1996). A longitudinal study of upward feedback. *Personnel Psychology, 49*, 703-740.

- Richardson II, R. F. (2010). 360-degree feedback: Integrating business know-how with social work values. *Administration in Social Work, 34*, 259-274.
- Roberts, J. M. (1995). The introductions to team-based organizations. *The History of the World*. (pp. 18-34). Harmondsworth, UK: Penguin.
- Robertson, I. T., & Kinder, A. (1993). Personality and job competences: The criterion-related validity of some personality variables. *Journal of Occupational and Organizational Psychology, 66*, 225-244.
- Ross, L. (1977). Social roles, social control, and biases in social perception process. *Journal of Personality and Social Psychology, 35*, 485-494.
- Rothwell, W. J. (2005). *Effective Succession Planning: Ensuring Leadership Continuity and Building Talent from Within*. New York, NY: American Management Association.
- Rushton, J. P., & Irwing, P. (2009). A general factor of personality in 16 sets of the Big Five, the Guilford-Zimmerman Temperament Survey, the California Psychological Inventory, and the Temperament and Character Inventory. *Personality and Individual Differences, 47*, 558-564.
- Schippmann, J. S., Ash, R. A., Battista, M., Carr, L., Eyde, L. D., Hesketh, B. et al. (2000). The practice of competency modeling. *Personnel Psychology, 53*, 703-740.
- Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods, 1*, 115-129.

- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 66*, 261-273.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 12*, 262-274.
- Schmidt, F. L., & Kaplan, L. B. (1971). Composite vs. multiple criteria: A review and resolution to the controversy. *Personnel Psychology, 24*, 419-434.
- Scullen, S. E. (1997). When ratings from one source have been averaged, but ratings from another source have not: Problems and solutions. *Journal of Applied Psychology, 82*, 880-888.
- Scullen, S. E., Mount, M. K., & Judge, T. A. (2003). Evidence of the construct validity of developmental ratings of managerial performance. *Journal of Applied Psychology, 88*, 50-66.
- Seifert, C. F., Yukl, G., & McDonald, R. A. (2003). Effects of multi-source feedback and a feedback facilitator on the influence behavior of managers toward subordinates. *Journal of Applied Psychology, 88*, 561-569.
- Silvester, J., & Wyatt, M. (2012). *Using 360-degree review to assess stakeholder perceptions of political leadership*. Paper presented at the annual conference of the Society of Industrial and Organizational Psychology, San Diego, CA.
- Smither, J. W., London, M., Flautt, R., Vargas, Y., & Kucine, I. (2002). *Does discussing multi source feedback with raters enhance performance improvement?* Paper presented at the annual conference of the Society of Industrial and Organizational Psychology, Toronto.

- Smither, J. W., London, M., Flautt, R., Vargas, Y., & Kucine, I. (2003). Can working with an executive coach improve multi-source feedback ratings over time? A quasi-experimental field study. *Personnel Psychology, 56*, 23-44.
- Smither, J. W., London, M., & Reilly, R. R. (2005). Does performance improve following multi-source feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology, 58*, 33-66.
- Smither, J. W., London, M., & Richmond, K. R. (2006). Relationships between leaders' personality and reactions to, and use of, multi-source feedback: A longitudinal study. *Group and Organization Management, 2*, 1-34.
- Spencer, L. M., & Spencer, S. M. (1993). *Competence at Work: Models for Superior Performance*. USA: John Wiley & Sons, Inc.
- Steenkamp, J-B. E. M., & Baumgartner, H. (2000). On the use of structural equation models for marketing modeling. *International Journal of Research in Marketing, 17*, 195-202.
- Strategic Success Model. (2003). Technical Manual. *Strategic Success Model 360-degree Feedback Technical Manual*. 1-21.
- Technical Manual. (2009). Technical Manual. *GZTS/DFOS Personality Inventory Technical Manual*. 1-105.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44*, 704-742.
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*, 25-29.

- Thornton, G. C., III. (1980). Psychometric properties of self-appraisals of job performance. *Personnel Psychology, 33*, 263-271.
- Toegel, G., & Conger, J. A. (2003). 360-degree assessment: Time for reinvention. *Academy of Management Learning and Education, 2*, 297-311.
- Tornow, W. W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management, 32*, 221-229.
- Ulrich, D. (1997). *Human Resource Champions*. Boston, MA: Harvard Business School Press.
- Ulrich, D., & Lake, D. (1990). *Organizational Capability*. New York: Wiley.
- Uniform Guidelines on Employee Selection Procedures. (1978). Federal Register, 43, 38290-38315.
- Vancouver, J. B., & Day, D. V. (2005). Industrial and organizational research on self-regulation: From constructs to applications. *Applied Psychology: An International Review, 54*, 155-185.
- Van Velsor, E. (1998). Designing 360-degree feedback to enhance involvement, self-determination, and commitment. In W. W. Tornow and M. London (Eds.), *Maximizing the Value of 360-degree Feedback: A Process for Successful Individual and Organizational Development* (pp. 321-344). San Francisco, CA: Jossey-Bass.
- Van Velsor, E., Taylor, S., & Leslie, J. (1993). An examination of the relationships among self-perception accuracy, self-awareness, gender, and leader effectiveness. *Human Resource Management, 32*, 249-264.
- Vazirani, N. (2011). Competencies and competency models: A brief overview of its development and application. *SIES Journal of Management, 7*, 121-131.

- Violato, C., Lockyer, J. M., & Fidler, H. (2008). Changes in performance: A 5-year longitudinal study of participants in a multi-source feedback program. *Medical Education, 42*, 1007-1013.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81*, 557-574.
- Waldman, D., Atwater, L., & Antonioni, D. (1998). Has 360-degree feedback gone amok? *Academy of Management Executive, 12*, 86-94.
- Walker, A. G., & Smither, J. W. (1999). A five-year study of upward feedback: What managers do with their results matters. *Personnel Psychology, 52*, 393-423.
- Warr, P. B. (1999). Logical and judgmental moderators of the criterion-related validity of personality scales. *Journal of Occupational and Organizational Psychology, 72*, 187-204.
- Warr, P. B. (2000). Indirect processes in criterion-related validity. *Journal of Organizational Behavior, 21*, 731-745.
- Warr, P. B., & Bourne, A. (1999). Factors influencing two types of incongruence in multi-rater judgments. *Human Performance, 12*, 183-210.
- Warr, P. B., & Hoare, S. (2002). Personality, gender, age and logical overlap in multi-source ratings. *International Journal of Selection and Assessment, 10*, 279-291.
- Weiss, H. M., & Adler, S. (1984). Personality and organizational behavior. In Staw, B. M., Cummings, L. L. (Eds.), *Research in Organizational Behavior*. (pp. 1-50). Greenwich, CT: JAI Press.

- Wohlers, A. J., & London, M. (1989). Ratings of managerial characteristics – evaluation difficulty, co-worker agreement and self-awareness. *Personnel Psychology*, *42*, 235-261.
- Wood, L., Hassell, A., Whitehouse, A., Bullock, A., & Wall, D. (2006). A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. *Medical Teacher*, *28*, 39-61.
- Yammarino, F., & Atwater, L. (1993). Understanding self-perception accuracy: Implications for human resource management. *Human Resource Management*, *32*, 231-247.
- Yammarino, F. J., & Atwater, L. E. (1997). Do managers see themselves as others see them? Implications of self-other rating agreement for human resource management. *Organizational Dynamic*, *25*, 35-44.
- Yukl, G., & Van Fleet, D. D. (1990). Theory and research on leadership in organizations. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (pp. 421-456). Palo Alto, CA: Consulting Psychologists Press.