

Florida Institute of Technology

Scholarship Repository @ Florida Tech

---

Theses and Dissertations

---

5-2023

## Exploring Factors that Influence Disagreements in Self-Other Ratings in 360-Degree Feedback Systems

Robert Allen Miller

Follow this and additional works at: <https://repository.fit.edu/etd>



Part of the Industrial and Organizational Psychology Commons

---

Exploring Factors that Influence Disagreements in Self-Other Ratings in 360-Degree Feedback Systems

by

Robert Allen Miller

A dissertation submitted to the School of Psychology of  
Florida Institute of Technology  
in partial fulfillment of the requirements  
for the degree of

Doctor of Philosophy  
in  
Industrial/Organizational Psychology

Melbourne, Florida  
May 2023

We, the undersigned committee, hereby approve the attached dissertation,  
“Exploring the Factors that Influence Disagreements in Self-Other Ratings in 360-  
Degree Feedback Systems.”

by  
Robert Allen Miller

---

Gary N. Burns, Ph.D.  
Professor  
School of Psychology  
Major Advisor

---

Jessica L. Wildman, Ph.D.  
Associate Professor  
School of Psychology

---

Patrick D. Converse, Ph.D.  
Professor  
School of Psychology

---

Kimberly N. Sloman, Ph.D., BCBA-D  
Associate Professor  
School of Behavior Analysis

---

Robert A. Taylor, Ph.D.  
Professor and Dean  
College of Psychology and Liberal Arts

# Abstract

Title: Exploring Factors that Influence Disagreements in Self-Other Ratings in 360-Degree Feedback Systems

Author: Robert Allen Miller

Major advisor: Gary Burns, Ph.D.

360-degree feedback systems are some of the most widely used ratings systems in organizations today. The current study examined the use of 360-degree feedback systems as a method for capturing job performance in organizations and the challenges associated with accuracy in ratings due to individual differences in perception. The study investigated the links between self-other rating tendencies and narrow facets of personality and examined how an individual's personality affects self- and other-ratings of their performance behaviors. The findings suggest that various aspects of personality differentially affect both self- and other-ratings of performance, providing a deeper understanding of how individual differences such as personality affect ratings of performance from multiple perspectives, including rating source, self-enhancement, self-verification, and self-presentation. The study highlights the importance of applying appropriate methods to study the prediction of self-other agreement and the need to revisit current theories of self-other agreement to consider the complex multivariate links of how personality interacts with perceptions of behavior.

# Table of Contents

Abstract .....	iii
List of Figures .....	vi
List of Tables.....	vii
Acknowledgement.....	viii
Dedication .....	ix
Chapter 1 Introduction .....	1
Chapter 2 Literature Review .....	7
Self-Other Ratings in 360-Degree feedback .....	7
Social Relations Model .....	18
Self-Perception Theory .....	20
Self-Enhancement Theory.....	20
Self-Presentation .....	21
Symbolic Interactionism .....	22
Self-Verification.....	22
SRM Summary.....	22
A Self-Other Rating Agreement Performance Appraisal Model .....	24
Biographical Characteristics .....	25
Cognitive Processes .....	26
Context Factors .....	28
Individual and Personality Characteristics.....	32
Dominance and Assertiveness .....	38
Sociability .....	39
Well Being .....	41
Empathy .....	43
Good Impression .....	44
Chapter 3 Method.....	46
Participants.....	46
Measures .....	47

California Personality Inventory (CPI) .....	47
Customized multi-rater instrument .....	50
Procedures .....	53
Chapter 4 Results .....	54
Preliminary Analyses .....	54
Hypothesis Testing.....	55
Chapter 5 Discussion .....	72
Theoretical Implications.....	75
Practical Implications.....	77
Limitations .....	77
Future Research Directions .....	79
References .....	82
Tables .....	104

## List of Figures

Figure 1: The Effect of Good Impression on Self and Peer Ratings.....	67
Figure 2: The Effect of Good Impression on Self and Other Ratings.....	67

## List of Tables

Table 1: CPI 260 Scales and Their Purposes .....	104
Table 2: 360-degree Instrument Scale Descriptions .....	107
Table 3: Personality, Performance, and Rating Tendency Correlations .....	108
Table 4: Multivariate Regression Results .....	109
Table 5: Exploratory Correlations.....	110



## Acknowledgment

I cannot fully express my gratitude for the support of my faculty, friends, and family. I would first like to convey my deepest gratitude to my advisor Dr. Gary Burns. His advice, insight, and encouragement helped me to grow both personally and professionally. I am also grateful for the support and encouragement of Dr. Cheryl Stenmark, my first professor in I-O Psychology, who opened my eyes to the field and helped me find something I could be passionate about. I am thankful for the faculty at Florida Tech and Angelo State University that helped me learn and grow. I would like to especially thank my close friend, the future Dr. Charlie Blömstrom-Johnson for his willingness to listen to my incessant complaining and for his “sage” advice. May it never steer me wrong.

I want to say thank you to my boss, friend, and mentor, Dr. Liam Ryan for his unwavering support through this process. He and my other co-workers at CMA Global, Inc. have provided me with the environment and tools to learn how to apply what I’ve learned in a meaningful way and I couldn’t be more grateful.

Finally, I would like to thank my parents, sister, and children for their unwavering love and support throughout this journey. Their words of encouragement and tolerance of my periodic absences from their lives will not be forgotten.

## Dedication

I would like to dedicate this work to my wonderful, awesome, brilliant, beautiful, caring, funny, and supportive wife, Rachel. I would never in a million years have expected to be so blessed by having you in my life. There is no way, in any dimension, that I would have made it here without you. You inspire me every day and push me to be a better man, husband, and father. I love you, and I'm so glad that I get to spend the rest of my life with you. I can't wait to see where it takes us.

# Chapter 1

## Introduction

One of the most critical questions facing organizations regarding employees is how to accurately capture job performance in such a manner that it is useful for both administrative and developmental purposes. For the past 30+ years, 360-degree feedback, also known as a multi-source or multi-rater assessment, has become an increasingly popular method for answering this question (Fletcher, 2014). As opposed to traditional one-way supervisor-generated ratings, 360-degree feedback systems gather further input from the employees themselves, their peers, and their direct reports to provide a more holistic picture of employee performance (Kim et al., 2016).

Part of the reason for using a more holistic approach is that “individuals have a significantly different view of their own job performance than that held by other people” (Thornton, 1980, p. 268). When employees receive feedback from supervisors with which they disagree, it can be attributed to the personal characteristics of the supervisor or other factors in a negative manner which can lead to lower job performance (Cawley et al., 1998). More feedback from more sources is one benefit that 360-degree feedback offers that can help employees receive feedback more productively (Facteau et al., 1998). Other benefits of using

360-degree feedback include better financial performance (Kim et al., 2016), improved performance (Smither et al., 1995), enhanced perceptions of the appraisal system (Cawley et al., 1998), and improved communication and employee satisfaction (Hall et al., 1996). While research has highlighted the benefits of using 360-degree feedback, the practice does have its challenges (Fletcher, 2014).

Nearly a century of research has found that the biggest problem with job performance appraisal is accuracy in ratings (Austin & Villanova, 1992). 360-degree feedback systems are no exception. Because these systems use multiple sources, each introducing their own bias, 360-degree feedback has even more challenges with accuracy than other systems (Viswesvaran et al., 1996). One aspect of these systems' accuracy problem is a significant lack of agreement among raters (Atwater et al., 2007). As Thornton (1980) pointed out in his quote above, subsequent research has found that the most common discrepancy in 360-degree feedback is a lack of agreement between self-ratings and ratings provided by supervisors, peers, and others (Brett & Atwater, 2001; Viswesvaran et al., 1996). Most research has explored this lack of agreement by framing it as measurement error in interrater reliability.

In terms of job performance appraisal, interrater reliability is defined as “the extent to which raters agree on the performance of ... individuals” (Viswesvaran et al., 1996). It is important to note that reliability in rating variability falls under source

disagreement, which refers to the differences between raters because each rater or rater group represents a unique perspective of the measured performance (Farr & Newman, 2001). For example, the direct reports of a person being assessed may disagree with the supervisor's ratings because they each view different aspects of the ratee's performance. Interrater reliability and source disagreement are conceptually distinct and always present. Together, they characterize self-other agreement (SOA) (Viswesvaran et al., 1996).

When asked to rate themselves, research has found that individuals tend to consistently rate themselves in a specific direction compared to how others rate them (Yammarino & Atwater, 1993). Atwater and Yammarino (1992) identified three categories of self-rater: over-estimators, under-estimators, and in-agreement raters. Over-raters tend to rate themselves higher and under-raters lower compared to the ratings of others. In-agreement raters tend to rate themselves similarly to others.

The categories of rating tendencies have been associated with several organizational outcomes, including leadership effectiveness, leader development, and managerial performance (Fleenor et al., 2010). Over-estimators tend to set unrealistic, high goals for themselves and their subordinates, are unlikely to accept feedback due to feelings of already achieving high performance, and tend to have poor supervisor-subordinate relationships (Brutus et al., 1999; Fleenor et al., 1996;

Yammarino & Atwater, 1993). Under-estimators may fail to take the initiative due to low feelings of self-worth, which result in lower promotion rates but employees who are more interested in self-development and improvement (Yammarino & Atwater, 1993). Additionally, in-agreement ratings are positively associated with job performance and leadership effectiveness (Atwater & Yammarino, 1992; Bass & Yammarino, 1991; Van Velsor et al., 1993). Given these critical outcomes to organizational success, it is important to further explore rating tendencies to understand their relationships with employee characteristics. Understanding these relationships and their antecedents can be vital in developing better selection, training, and development programs for organizations (Atwater & Yammarino, 1997; Yammarino & Atwater, 1993). Therefore, the current study explored factors that influence disagreements in rating tendencies between self and others.

In the following literature review, I will briefly review the uses of self-other ratings in 360-degree feedback systems. Then, I will provide an overview of the conceptual and theoretical foundations of self-perception and person perception using the Social Relations Model, which includes theories relating to self-verification, self-perception, self-presentation, and symbolic interactionism. These concepts will then be applied to a model of SOA as posited by Atwater and Yammarino (1997), which outlines SOA as an interaction of five categories: biographical characteristics, cognitive processes, contextual factors, job-relevant experiences, and individual personality characteristics. This model's factors outline aspects of

the self-rater and other raters that impact SOA. The current study focused on the five factors in relation to self-rater only. Other-rater aspects were not addressed in the current study because individual differences across multiple raters are likely to average out, allowing the focus to shift to the self-rater, which may offer a unique perspective.

Finally, I will present the hypotheses for the current study by concentrating on self-rated, work-related personality variables as an antecedent and explanation of SOA directionality. The current study explored directionality using ratings from participants, their supervisors, peers, and direct reports. As previously discussed, the perspectives and ratings from each group likely differ. As such, each group was examined separately.

The current study examined the relationship between the self-rater's work personality dimensions and their rating tendencies on job performance. Beyond exploring these relationships, the current study attempted to predict rating tendencies based on scores on a commonly used personality inventory. Analysis of the gathered data was intended to yield insights that will likely be useful for increasing the effectiveness of 360-degree feedback for individual and organizational development as well as potentially adding incremental validity to selection systems using work personality assessments. Understanding these relationships can improve interventions designed to increase receptivity to feedback

which can improve performance (Harris & Schaubroeck, 1988). Further, understanding these relationships can provide insight into training and development programs, gearing them toward specific personality characteristics and rating tendencies. Finally, the findings will provide a deeper understanding of the theoretical relationship between personality, self-perception, and self-awareness.



## Chapter 2

### Literature Review

#### Self-Other Ratings in 360-Degree feedback

Despite recent trends to eliminate performance ratings in the workplace, they remain the most widely used method of measuring job performance (Evans-Reber, 2020). In general, performance ratings are used for various reasons, including high-potential identification, employee development, promotions, and succession planning. 360-degree feedback systems are one form of performance rating system that is becoming more prevalent in the workplace. In the late 1990s, it was estimated that 25% of organizations used these types of rating systems (Antonioni, 1996). That estimate has since risen to more than 50% of companies using 360-degree feedback systems for administrative purposes and more than 70% of top companies using them for developmental purposes (Church et al., 2019; Church & Rotolo, 2013).

Before discussing the benefits and challenges of 360-degree feedback, it is essential to characterize these systems. A recent review by Bracken et al. (2016) defined 360-degree feedback as:

360° Feedback is a process for collecting, quantifying, and reporting coworker observations about an individual (i.e., a ratee) that facilitates/enables three specific data-driven/based outcomes: (a) the collection of rater perceptions of the degree to which specific behaviors are exhibited; (b) the analysis of meaningful comparisons of rater perceptions across multiple ratees, between specific groups of raters for an individual ratee, and for ratee changes over time; and (c) the creation of sustainable individual, group, and/or organizational changes in behaviors valued by the organization. (p. 764)

The rise in the prevalence of 360-degree feedback can be attributed to many factors. Compared to traditional top-down performance ratings, 360-degree feedback systems provide a more comprehensive view of individual job performance because the observational data is gathered from all key stakeholders in the workplace (Church et al., 2019). This process benefits organizations in a few ways. First, employee performance is assessed from multiple points of view as opposed to a single view by a single supervisor. Generally, supervisors are more likely to observe performance outcomes primarily. Peers and direct reports provide perspectives that are likely not observed by supervisors alone and can impart a more complete understanding of day-to-day employee behaviors such as time management and counter-productive work behaviors (Harris & Schaubroeck, 1988).

Another way that 360s benefit organizations are through their effects on employees and their performance. Employees are more likely to feel accountable to peers, customers, and subordinates when they know that their input will be included in the employee's performance appraisal (Bracken et al., 2016; Ward, 1997). This feeling of accountability can lead to better working relationships and higher performance (Roethlisberger & Dickson, 2003). Additionally, when employees perceive feedback they receive as being from multiple sources, they are more likely to give it more credence, especially when the feedback is for developmental purposes (Kluger & DeNisi, 1996; Smither et al., 2005).

In the early 2000s, 360-degree feedback systems garnered much debate about the accuracy and validity of their ability to measure job performance, prompting calls for the practice to be used for developmental purposes only (Bracken et al., 2001; London & Smither, 2002). Since then, its popularity and research advances have made the practice a modern-day staple as more and more companies incorporate it into their performance management systems and use the results in organizational decision-making (Church & Rotolo, 2013). However, there is still much to be learned about the accuracy and validity of 360-degree feedback.

One of the most challenging aspects of 360-degree feedback is rooted in what makes it different from other performance rating systems. While 360s provide different perspectives of employee performance, those perspectives do not always

agree. In examining relationships between raters within the same rating group (i.e., supervisors, peers, direct reports), one meta-analytic study found the average correlation between ratings by two different supervisors to be .50, two peers averaged at .37, and two subordinates at only .30 (Conway & Huffcutt, 1997). Another study found correlations among ratings of .52 for managers and .46 for peers (Viswesvaran et al., 1996). This lack of agreement is not limited to raters in the same rating group. In fact, the agreement between groups shares a similar lack of consistency. Conway and Huffcutt (1997) found that supervisors and peers of the ratee presented a corrected agreement correlation of .79 in their ratings, supervisors and subordinates at .57, and peers and subordinates at .66. While these correlations highlight some of the potential validity of 360s, the real challenge concerning accuracy among rating groups is self-other agreement.

A large body of evidence exists cataloging the discrepancy between how employees view themselves and how others view them (Ostroff et al., 2004; Park & Judd, 1989). That discrepancy extends to employee performance and behavior in the workplace (Harris & Schaubroeck, 1988; Mabe & West, 1982; Smither et al., 2005). Mabe and West (1982) found a correlation of only .04 when comparing ratings between managers' self-ratings and ratings provided by others. Another meta-analytic study found corrected correlations of .36 and .35 for self-supervisor and self-peer ratings, respectively (Harris & Schaubroeck, 1988). These results lend

weight to Thornton's (1980) assertion that employees view their performance differently than others.

Several explanations for self-other rating disagreement have been posited in the literature. Harris and Schaubroeck (1988) sorted explanations into three categories: egocentric bias, differences in organizational level, and observational opportunities. Egocentric bias refers to the idea that self-ratings are biased in some way while other raters share a common perspective. This bias could be caused by rater defensiveness, where the self-rater is motivated by some contextual or personal factor to inflate their rating and thereby enhance their evaluation. For example, ratees with high self-esteem may be biased toward higher self-ratings, whereas ratees with low self-esteem may not. Another possible egocentric bias is posited by attribution theory, where ratees attribute good performance to themselves and poor performance to environmental factors (Kelley & Michela, 1980). Differences in organizational level refer to the concept that people at different levels of an organization have different perceptions of performance. Two versions of this concept have been posited. Some researchers suggest that raters at different levels weight performance differently, while other scholars maintain that raters at different levels define and measure performance differently (Harris & Schaubroeck, 1988). Observational opportunities are an explanation of disagreement that focuses on how raters have differing abilities or opportunities to observe rated behavior (Harris & Schaubroeck, 1988). For example, peers are posited to have more

opportunities to observe performance than supervisors. Varying levels of opportunity to observe can lead to differing assessments of appropriate ratings.

A more recent literature review outlined possible causes of disagreement into three categories: biographical differences, personality and individual characteristics, and job-relevant experiences (Fleenor et al., 2010). Biographical differences refer to characteristics such as age, gender, ethnicity, education levels, or level of position in an organization. For example, males and older raters tend to overrate their leadership effectiveness and abilities compared to other raters (Brutus et al., 1999; Visser et al., 2008). Ostroff et al. (2004) found that individuals with less education and non-whites displayed a tendency to overrate their performance compared to others. Other studies found that individuals in higher positions have displayed a tendency to rate themselves higher than other raters, likely due to a lack of appropriate feedback (Brutus et al., 1999; Gentry et al., 2007; Sala, 2003).

Personality and individual characteristics that could explain disagreements include traits such as agreeableness, extraversion, openness to experience, conscientiousness, narcissism, and dominance which were all found to be positively related to self-ratings of leadership (Fleenor et al., 2010; Judge et al., 2006). Neuroticism was found to be negatively related to self-ratings, and meta-cognitive ability and intelligence were found to have a positive relationship with in-agreement ratings with others (Furnham et al., 2005; Kruger & Dunning, 1999).

Finally, job-relevant experiences refer to the feedback frequency and quality to which the self-rater is exposed (Fleenor et al., 2010). In a two-year longitudinal study, Bailey and Fletcher (2002) found higher congruence in self and supervisor ratings when appropriate feedback processes were utilized. However, it should be noted that self- and supervisor- ratings were not analyzed separately; therefore, the congruence could have been due to changes in the ratings of the supervisor, the self-ratings, or both (Fleenor et al., 2010).

Many scholars urge caution when using self-appraisals in performance evaluation processes due to questions of accuracy and lack of agreement. Accuracy in this context can refer to either the accuracy of the self-ratings or other ratings, which constitute a lack of agreement (J.-L. Farh & Dobbins, 1989). Kenny (1994) points out that researchers should focus on consensus as a natural extension of accuracy. The author posits that although consensus does not equate to accuracy, if raters are accurate, then consensus will follow. Well-documented problems inherent to setting and measuring job performance criteria preclude rigorous research on accuracy, which has forced researchers examining SOA to focus on consensus (Austin & Villanova, 1992; Sinha, 2004). Other scholars argue that, when comparing self and other ratings, one should not presume that other-ratings are “true” scores of performance (Atwater & Yammarino, 1997). As early as the 1920s, research has found that self-ratings are problematic (Hoffman, 1923). Numerous studies have found self-ratings to be unreliable, inflated, biased, and inaccurate

(Harris & Schaubroeck, 1988; Mabe & West, 1982; Thornton, 1980). However, other studies have found that self-ratings do not consistently differ from other ratings (Fleenor et al., 1996, 2010). Dunnette (1993) concluded that self-ratings contained an element of accuracy and, therefore, others' ratings should not be "the capstone of accuracy against which self-descriptions [ratings] should always be validated" (p. 376). These findings have generated a stream of research focused on understanding factors that affect when self-ratees and others disagree and in which direction.

Early research examining SOA identified three categories of rating agreement: under-raters, over-raters, and in-agreement raters (Atwater & Yammarino, 1992; Van Velsor et al., 1993). Under-raters rate their performance lower than others rate them. In contrast, over-raters tend to rate their performance better than other-raters. These two rating groups are associated with several organizational outcomes, such as leadership effectiveness and job performance (Atwater & Yammarino, 1992; Van Velsor et al., 1993; Yammarino & Atwater, 1993). Additionally, agreement among self and peers is tied to leadership success (Halverson et al., 2002).

Later, Atwater and Yammarino (1997) proposed an SOA model that expanded their rating agreement categories to four, replacing the in-agreement category with in-agreement/good and in-agreement/poor. In-agreement/good raters refer to individuals whose self-ratings are high and similar to ratings given by others.



Conversely, in-agreement/poor raters are those individuals whose self-ratings are poor and agree with other-ratings. The authors posit that ratings that agree, whether they are good or poor performers, are better than those that disagree because agreement indicates a shared view of performance (Atwater & Yammarino, 1997). For example, if the ratee and the supervisor agree that the ratee's performance is low, that agreement provides a basis to start corrective interventions. If the rater and ratee cannot agree, such interventions would likely be less effective. Fleener et al. (1996) added a further two categories to the SOA model, breaking down over-estimators and under-estimators, adding both good and poor versions of each. Fleener and associates then found that these additional categories were unnecessary as under-raters did not typically fall into the under-rater/poor category, and over-raters did not significantly fall into the over-rater/good category (Fleener et al., 1996).

Research has found relationships between the rating tendency and important individual and organizational outcomes. At the individual level, in-agreement raters have been associated with higher performance compared to over- and under-raters (Atwater & Yammarino, 1992). Van Velsor et al. (1993) reported that over-raters garnered lower ratings from subordinates than in-agreement and under-raters. Higher levels of agreement were also related to performance, organizational level, and compensation (Ostroff et al., 2004). Bass and Yammarino (1991) also found that self-rating accuracy was positively related to managers' promotability. At the

organizational level, Atwater and Yammarino (1992) found that promotion endorsements from supervisors were positively related to effectiveness for in-agreement raters and negatively related for over-raters. In a study of 83 supervisor-subordinate dyads, Szell and Henderson (1997) observed that agreement was positively related to subordinate organizational commitment and job satisfaction.

Several behavioral patterns were found by Yammarino and Atwater (1997) relating to the type of rating tendency. Under-estimators and over-estimators were shown to be inaccurate in identifying their strengths and weaknesses and, as a result, made poor job-related decisions. Under-raters were found to be generally successful, effective, and pleasant to be around despite their emotional highs and lows. On the other hand, over-raters showed tendencies to hold negative attitudes such as resentment and hostility and were more likely to suffer from career derailment (Yammarino & Atwater, 1997). Their research provided substantive evidence that in-agreement raters rated high in performance will most likely be the best performers, have positive job attitudes, get along well with coworkers, and have low turnover and absenteeism. Also, they are more likely to make effective decisions and alter their behavior based on constructive feedback from others. In contrast, in-agreement raters who are rated poor performers are most likely to be low performers, have low organizational commitment, and make poor job-related decisions. These individuals are also likely to have higher absenteeism and turnover, negative job attitudes, and lower self-worth (Yammarino & Atwater,

1997). The silver lining for in-agreement (poor) raters is that they are more likely to be accurate in their self-assessment, which could be a starting point for interventions if they are willing to engage in them.

Due to the evidence of linkages between rating tendencies and various important individual and organizational outcomes, researchers have studied possible antecedents associated with overrating, under-rating, and in-agreement rating inclinations. Cognitive processing is related to rating tendencies (J.-L. Farh & Dobbins, 1989). Farh and Dobbins (1989) suggested that their findings were due to inherent differences in psychological processes between observers (other-raters) and actors (self-raters), as posited in attribution theory. Other researchers suggest that raters have unique perspectives on different aspects of job performance and that the rating agreement is caused by more contextual factors (Klimoski & London, 1974). At the individual level, self-raters who provide ratings consistent with other raters have been found to have higher cognitive ability, greater memory capacity, and greater ability to process information (Yammarino & Atwater, 1997). Atwater and Yammarino (1997), in their seminal model of self-other agreement, posited that individuals with a more stable mental and emotional state are more likely to provide ratings that are in-agreement with others. Many studies have examined personality as an antecedent of rating agreement; however, most of these studies have focused on the Big 5 personality traits. There is currently a dearth of research examining other models of personality, such as the California Personality

Inventory (CPI). Considering that the CPI is a widely used personality measurement in business and that organizational selection and development systems can benefit from a greater understanding of how personality, as measured by the CPI, can impact important outcomes, it follows that further investigation of personality factors that can predispose one to rate one way or the other is warranted. Before exploring the role of personality in self-rater rating tendencies, two streams of theory shall be examined. First, I will delve into the social relations model to consider self-perception theory and discrepancies that occur between self and other perceptions. Next, I will apply that model to the literature covering self-other agreement in performance appraisal. These topics will provide a theoretical foundation and framework for hypothesizing the effects of personality in self-other rating agreement.

## Social Relations Model

Similarities and differences in self-other ratings result from how a person views themselves, how they view others, and the interaction between themselves and others (Kenny, 1994). These concepts are theorized as self-perception, person perception, and interpersonal perception, respectively. While these theories have extensive threads of research in and of themselves, the Social Relations Model (SRM) is useful for organizing them into a framework that is useful in examining self-other agreement. SRM focuses on the self, the other, and the interaction

between the two in defining how our self-perception and other perception interact to set the stage for our perceived reality in terms of our interactions with others. SRM posits that interpersonal perception is comprised of four components (Kenny, 1994). Actor, the perceiver, and partner, the target of those perceptions, are the two main components of interpersonal perception. The perceiver effect refers to how the actor perceives the partner or target, and the target effect refers to how others generally see the partner or target. Relationship effect is a third component that refers to how a person sees any given target. While self-perception theory focuses on how one views the self and person perception theory focuses on how one views others, SRM focuses on the variances in the relationship effect between the actor and the target while controlling for perceiver and target effects (Kenny, 1994). The fourth effect is the constant effect which refers to the average across all ratings of perceivers, targets, and relationships.

The SRM outlines five theories that endeavor to explain how an individual's self-perception impacts their behaviors and views. These theories are, as summarized by Kenny (1994), self-perception theory ("I see myself as I behave"), self-enhancement theory ("I see myself as better than others"), self-presentation theory ("I present myself to others as they want to see me"), symbolic interactionism ("I see myself as significant others see me"), and self-verification theory ("I make others see me as I see myself") (p. 180). In a given situation, it is possible that none of these theories explain an individual's viewpoint entirely or may even fail to

touch on it. They all likely have some validity in general when examining self-other perception.

### *Self-Perception Theory*

The study of self-perception in psychology dates back as far as William James in the late 1800s and has developed into several different schools of thought. Bem (1967) posited that self-perception occurs when one views their behaviors in the same way that they view others to determine what traits, attitudes, beliefs, feelings, and thoughts one has. However, Fenigstein et al. (1975) proposed that people are introspective and cognitively consider perceptions of their own characteristics that lead to their behaviors. They also proposed that the level of introspection and self-consciousness differs between people and that individuals vary by focusing on what they perceive about themselves and what they fear others perceive about them. Presumably, those higher in self-consciousness and introspection are less likely to use Bem's method of self-perception (Kenny, 1994). Both theories imply that behavior can be modified based on personal beliefs leading to a more significant potential for agreement in self and other perceptions when based on the same observable behavior.

### *Self-Enhancement Theory*

Self-enhancement theory is a theory of self-perception wherein actors focus more on their positive characteristics and downplay or ignore their negative ones (Kenny,

1994). This process is engaged to make one feel better about themselves, and although it is a biased form of information processing, it is inherent to human nature (Kenny, 1994). Some have even posited that self-enhancement is required for mental health (Taylor & Brown, 1988). Through distorting and filtering information, individuals can more easily maintain self-efficacy and a positive view of their future capabilities. Therefore, self-ratings will contain some degree of self-enhancement. However, self-enhancement's focus on how one views the self is a discrete concept from self-presentation, which focuses on the perceived views of others.

### *Self-Presentation*

Another view of self-perception is that actors view themselves in relation to how they want other actors to view them (Kenny, 1994; Lewis & Neighbors, 2005). Otherwise known as impression management or social desirability, self-presentation varies between individuals (Crowne & Marlowe, 1960). This view implies that individuals change their behavior to match their perceptions of what they think will result in their desired outcome in terms of how someone else perceives them (Snyder, 1974). Concerning performance management, self-enhancement theory implies that individuals will change their behavior to match what they believe will result in good ratings. This will likely result in higher self-ratings in an attempt to present the best view of themselves possible.

### *Symbolic Interactionism*

Symbolic interactionism is like self-enhancement in that the actor is aware of and is impacted by the views of others. However, in this theory, actors sometimes see themselves the same way they perceive that significant others see them (Carter & Fuller, 2015; Kenny, 1994). Presumably, those who engage in symbolic interactionism are more likely to be in rating agreement with others.

### *Self-Verification*

The last theory that explains self-perception in the SRM is self-verification. This theory also takes the views of others into account. However, actors do so only to verify that others see them as they see themselves (Kenny, 1994; Smith et al., 2022). Self-verification occurs when actors attempt to convince others instead of being influenced by them. Interestingly, actors attempt to preserve or reinforce their sense of self by verifying their views through others. Actors who engage in self-verification are more likely to be prone to confirmation bias as the more strongly they believe that they have a given characteristic or personality trait, the more likely they are to project that view to others, ignoring any feedback that states otherwise (Swann et al., 1989).

### *SRM Summary*

With these five theories, the SRM attempts to outline how an individual's perception of themselves impacts self-other agreement in the context of



interpersonal perception and social relations. However, the SRM has some notable limitations. First, the SRM focuses on acquaintance, observability of traits, and evaluative extremity as the only potential influences of self-other agreement (Kenny, 1994). Acquaintance refers to how well a rater knows the target in terms of length of association and depth of relationship and is linked with self-other agreement (McCrae, 1982). Observability of traits refers to how easily traits can be perceived. Some traits, such as extraversion, are more visible and consistent than others and may contribute more to self-other agreement (Funder & Dobroth, 1987). Finally, evaluative extremity refers to how the perceiver feels about the observed trait in the target. Raters have been shown to overrate traits for which they have strong negative feelings (John & Robins, 1993). While these studies provide some support that the three listed factors have a relationship with self-other agreement, they focused solely on the characteristics of the target, and the traits of the raters were not considered. Another limitation is that the studies in question relied upon simple correlations that do not provide a clear picture of what caused the agreement, or lack thereof, between the perceiver and the target. Finally, SRM focuses on personality traits evidenced by behavior but does not explore how self-perception informs self-other agreement in job performance appraisal.

While social psychology has examined self-perception and how it influences behavior, personality traits and their effect on self-other agreement have also been explored in other contexts. Within the field of personality psychology, the social-

cognitive paradigm focuses on how individual characteristics impact the interaction of the individual with their environment (Funder, 2006). Personality traits have also been examined in the performance appraisal literature as possible antecedents to self-other agreement while examining their relationship with outcomes at the organizational level.

## A Self-Other Rating Agreement Performance Appraisal Model

In their seminal work on self-other agreement, Atwater and Yammarino (1997) outline a model focusing on organizing theories of agreement around the concept of rating accuracy. Accuracy is defined in their model as, “some criterion measure which may or may not itself be completely accurate” (Atwater & Yammarino, 1997, p. 124). The current study concentrates on the criterion of performance ratings rendered by individuals and supervisors, peers, and direct reports. As postulated by symbolic interactionism, accuracy in both self- and other-ratings inevitably influences agreement between the ratings. Atwater and Yammarino (1997) state,

In the ideal case, both self and other ratings are unbiased, informed, accurate, and the self-other rating comparison process will result in ratings that are in-agreement. However, in many instances, because ratings are biased and/or uninformed, self-others rating comparisons do not result in-

agreement. Rather, self-ratings may be (much) higher or lower than other ratings. (1997, p. 126)

Their self-other agreement model identifies possible antecedents to accuracy in self-other agreement, including biographical characteristics, cognitive processes, contextual or situational factors, and individual personality characteristics (Atwater & Yammarino, 1997). Research in these four areas in the context of self-other rating agreement is reviewed below.

### *Biographical Characteristics*

An array of biographical characteristics have been examined for their relationship with self-ratings, including age, gender, position in the organization, and tenure. Ostroff et al. (2004), in a study across 527 organizations, found that older managers tended to rate themselves higher in leadership compared to younger managers, while they were rated lower by their subordinates than their younger counterparts. These findings support a number of other studies that indicate a positive relationship between age and over-rating (Brutus et al., 1999; Moshavl et al., 2003; Vecchio & Anderson, 2009). With regard to gender, several studies have shown that females tend to rate themselves less positively than males, are more open to feedback, and are more accurate in their ratings (McKee et al., 2018; Vecchio & Anderson, 2009; Visser et al., 2008; Wohlers & London, 1989). As for position in the organization, Gentry et al. (2007) replicated the results of previous research

(Ostroff et al., 2004; Sala, 2003) by demonstrating that higher managerial positions indicated higher disagreement between self- and other-ratings. The authors suggested a lack of effective feedback likely contributed to this gap (Gentry et al., 2007). Finally, the relationship between tenure and self-other agreement has largely been found to be positive. Rothstein (1990), using a sample of nearly 10,000 managers across organizations, found that, as tenure increased, so too did rating agreement. However, Brief et al. (1977) discovered opposing findings in a study with 129 food services employees wherein the longer the tenure, the less agreement existed between supervisor and self-ratings. The authors noted that these results were unexpected and suggested that they may result from an artifact of that particular data set.

### *Cognitive Processes*

Atwater and Yammarino (1997) discussed two groupings of cognitive processes that they posited to have an impact on self-ratings. Schemas, beliefs, and expectations comprise the first group, while the second group comprises information and feedback. Schemas refer to “cognitive structure that helps organize existing and new knowledge” (Needham & Jacobson, 2020, p. 363). In forming a schema, assumptions and categories are created to organize and explain information that then affects memory, perception, and inferences. It follows that these assumptions will have an impact on how individuals perceive and rate

themselves and others. Attribution theory is an example of how assumptions impact ratings and how self- and other-ratings can disagree. This theory postulates that “actors (i.e., self-raters) attribute good performance to their behavior and poor performance to environmental factors” (Harris & Schaubroeck, 1988, p. 45; Kelley & Michela, 1980). However, other raters attribute success and failure differently, depending on how well they “like” the target. Success from a liked individual and failure from a disliked one is attributed to personal factors, whereas opposite or unexpected outcomes are attributed to situational factors beyond the control of the target (Kelley & Michela, 1980). Another cognitive process in this first grouping focuses on the beliefs an individual holds about themselves and others regarding success and failure. These beliefs can significantly impact self-ratings (Ashford, 1989). For example, if an individual feels that their failure was due to factors out of their control, they are less likely to rate themselves lower on performance than if they believe that the failure was due to some fault of their own (Ashford, 1989). Expectations have an impact on self-ratings as well. Individuals with low self-efficacy have been shown to under-rate their performance, while those high on this scale tend to overrate (Fleenor et al., 2010). Levy (1993) found that locus of control was significantly related to self-ratings. Participants in his study with a higher internal locus of control tended to rate themselves higher in managerial potential. However, it should be noted that locus of control in general and the managerial potential measured in this study are future-focused and may have little relation to

past-focused performance appraisals. All of these cognitive processes in the first grouping are also likely impacted by individual differences in self-awareness and receptivity to feedback (Sinha, 2004).

How individuals use information and feedback comprise the second group of cognitive processes posited by Atwater and Yammarino (1997) to influence ratings. Individuals who are receptive to feedback are also likely to allow that feedback to impact their self-evaluation and ratings (Bailey & Fletcher, 2002). Whether feedback is given formally or informally, the credibility of the person giving the feedback plays a role in how much that feedback can influence self-ratings. Additionally, if the feedback includes new, positive, specific, and consistent information, it is more likely to be heard by the individual receiving the feedback (Kinicki et al., 2004).

### *Context Factors*

Factors within the rating environment or context, such as environmental pressures, comparative information, specificity of information, rater experience, and political influence, have all been postulated to affect self- and other ratings (Atwater & Yammarino, 1997). Knowing that ratings can or will be compared to an objective criterion has been found to increase the accuracy of self-ratings and is an example of contextual pressure that may influence actors (Farh & Werbel, 1986). Another example is the tendency to over-rate when the evaluation is known to be for

administrative evaluative purposes rather than developmental (J. Farh & Werbel, 1986). Accountability, or the pressure to justify ratings, has also been found to enhance accuracy in those ratings (Tetlock, 1983; Tetlock & Boettger, 1989). The amount of comparative information can affect ratings in that they tend to be more distorted when less information about the measured criterion is available and more accurate when more information is presented (Atwater & Yammarino, 1997). Comparative information in performance appraisal refers to information about how others are performing, enabling individuals to compare their performance to that of others (Atwater & Yammarino, 1997). Farh and Dobbins (1989), for example, found higher correlations between self-ratings and objective performance indicators when rates were presented with comparative information about their peers' work than when not. Another aspect of information that has an effect is the specificity of the available information. Vague definitions of expected performance criteria can lead to distorted or inflated ratings where clearly defined objectives can lead to more accuracy and, therefore, agreement between raters (Ashford, 1989; Wohlers & London, 1989). Another influence on self-rating accuracy is the amount of experience individuals have in rating their abilities and behaviors. Research suggests that self-rating accuracy increases as raters garner experience in evaluating their own actions (Levine et al., 1977; Mabe & West, 1982). Atwater and Yammarino (1997) posit that this effect should be more prominent when raters can compare their self-ratings with other criteria, such as other-ratings or an objective

standard. The final contextual factor discussed in the model is political influence which can provide a wide range of situational influences on self- and other-ratings (Ferris & Judge, 1991). For example, supervisors may rate employees higher if more accurate ratings could damage their relationship with a subordinate or subordinates may rate themselves lower to signal humility and willingness to grow (Fleenor et al., 2010). Supervisors adjusting their ratings to be comparable to the ratings that other supervisors give subordinates to signal relative perceived standing is another example of political influences on ratings.

While demographics, cognitive processes, and contextual factors have been shown to influence self-other agreement, individual differences, and personality characteristics also play a vital role in how individuals perceive and evaluate themselves and others. In fact, examination of any of these four areas will inevitably include discussions of personality traits. For example, in biographical factors, it has been found that women and men differ in the self-ratings they report. Specifically, women were found to provide more accurate self-ratings than men because of a smaller degree of over-estimation. A possible explanation of these findings is based on possible personality differences between the genders, such as women scoring lower on self-esteem than men, a trait associated positively with self-ratings (Feingold, 1994; Goffin & Anderson, 2007). Self-esteem is also discussed in examinations of cognitive processes. Individuals with higher self-esteem are less impacted by feedback than those with lower self-esteem, which



would then influence self-ratings (Wood et al., 1994). Individuals with lower self-esteem are more likely to modify their self-evaluations based on how they perceive others perceive them, which would be exhibited by lower self-ratings. Thus, the personality characteristic of self-esteem would blend with cognitive processes to impact self-other agreement. This same effect relates to context factors. When in a low feedback environment, low self-esteem and moderate to high levels of neuroticism can reduce accuracy in self-ratings, while high self-esteem can mitigate some of the potential benefits of a high feedback environment due to lower feedback acceptance (Fleenor et al., 2010; Wood et al., 1994). While self-esteem is one example of personality characteristics that can influence self-other agreement, it is logical to assume that other personality traits can play a role and therefore merit further investigation as possible antecedents to self-other agreement, as posited by Atwater and Yammarino (1997). Furthermore, advances in analytical best practices have allowed a resurgence in in-depth analyses of self-other agreement as it relates to personality and other factors. Finally, while several studies have examined self-other agreement in relation to personality characteristics (e.g., Antonioni & Park, 2001; Bergner et al., 2016; McKee et al., 2018; Sinha et al., 2012), no studies exist looking at the relationship between self-other agreement and personality, as measured by the CPI.

### *Individual and Personality Characteristics*

In their model on self-other agreement, Atwater and Yammarino (1997), based on findings at the time, proposed several individual characteristics as antecedents to self-rating accuracy, including intelligence, cognitive complexity, achievement status, and ability along with personality traits such as locus of control, self-monitoring, and self-esteem. Based on their findings, Campbell and Lee (1988) suggested that personality characteristics underlie self-rating accuracy when contrasted with other ratings. Their findings also suggest that rating tendencies, similar to personality, tend to be stable over time. In this context, the current research aims to continue to examine whether a “personality constellation” potentially explains rating tendencies, as suggested by previous research (Atwater et al., 1998).

A well-known phenomenon, the Dunning-Kruger Effect, highlights how intelligence or meta-cognitive ability can impact self-evaluation. Kruger and Dunning (1999) found that an individual who scored low on an ability are more likely to overrate themselves on that ability stating, “unskilled individuals suffer a dual burden: Not only do they perform poorly but they fail to realize it” (p. 1131). However, London (1994) suggested that people with higher intelligence or cognitive complexity would be more likely to rate themselves more accurately because they can process and remember more accurate, relevant information. Beehr

et al. (2001) found that analytical, numerical, and cognitive abilities were all negatively related to performance self-ratings suggesting support for London's assertion.

Ability was also proposed as having an impact on self-ratings. Ruble and Flett (1988) found that elementary grade school children who scored lower in ability were less likely to seek self-evaluative information than those who scored higher, perhaps because they feared the information would be unflattering or negative. On the other hand, individuals with feelings of uncertainty regarding their abilities were more likely to show interest in evaluative information (Sorrentino & Hewitt, 1984; Trope, 1986). Atwater and Yammarino (1997) proposed that high achievement status would be related to higher accuracy in self-evaluations of ability based on several studies of a positive relationship between achievement status and self and peer rating accuracy (see Bailey & Lazar, 1976; Bayroff et al., 1954). Together, intelligence, achievement status, and ability have been inferred to denote capability, wherein individuals with higher capability can make more accurate judgments of their own and others' performance (Ashford, 1989). Ashford (1989) suggested that individuals with higher academic capability would be more likely to seek feedback for performance improvement resulting in higher awareness of performance standards and norms and, therefore, better rating accuracy. However, it is essential to note that the findings related to ability and achievement

status have only been substantiated in academic settings and have not been replicated in work contexts.

Another possible influence on self-other agreement is the personal characteristic of locus of control (Atwater & Yammarino, 1997). Using a hypothetical test of managerial performance, Levy (1993) found evidence of a positive relationship between internal locus of control and self-ratings. The author couched the findings in an attributional framework positing that individuals who believe they have more control over their performance are more likely to rate themselves higher. However, no causal directionality has been substantiated between locus of control and self-ratings. It may be that having an internal locus of control results in better performance, and therefore better ratings are founded in that as opposed to being based solely on the perception of control over their performance (Fleenor et al., 2010).

Self-monitoring refers to the ability to adapt one's behavior based on social cues from the environment (Day et al., 2002). Individuals high in self-monitoring would be expected to be more aware of how they are perceived and would use that information to rate themselves more accurately. Contrary to that expectation, however, high self-monitors have been found to rate themselves higher than others, and low self-monitors displayed greater self-other rating agreement (Miller & Cardy, 2000; Tunnell, 1980). The findings were explained as high self-monitors

showing lower consistency in behavior, leading to unreliable ratings, both for the self and from others. Low self-monitors were better able to describe their behavior and rate it appropriately due to more consistent behavior, which also led to more accurate ratings by others (Miller & Cardy, 2000).

Self-esteem has long been associated with an individual's self-evaluation. Self-esteem is defined as the evaluative component of self-concept (Brook, 1991). This concept has been extensively examined in relation to an individual's self-evaluation accuracy (Atwater & Yammarino, 1997; Fleenor et al., 2010). Two disparate theoretical models offer contrasting outcomes. Self-enhancement theory posited that people with low self-esteem would misrepresent their performance with higher ratings to compensate for their low self-worth (Schlenker et al., 1990).

Alternatively, self-consistency theory posited that people want their ratings to match their self-concept, which would result in accurate ratings (Wells & Sweeney, 1986). However, subsequent research has shown that individuals with high self-esteem are less likely to change how they perceive their behaviors based on feedback from others and are more likely to provide over-ratings than those with lower self-esteem (Kernis et al., 1991; Levy, 1993; Wood et al., 1994).

The social relations and self-other rating agreement models offer context for the extant research explaining rating discrepancies. In both models, individual differences have been shown to influence self-perception and the degree of other

rating agreement. Although individual differences such as self-esteem and locus of control have been extensively researched, other characteristics that could potentially influence rating tendencies have yet to be fully explored.

While several studies have examined the link between personality and self-other agreement, the preponderance of those studies have focused on the Big 5 and other personality traits as measured by instruments such as the IPIP, NEO, and HEXACO. To date, very little research has examined the link between self-other agreement and personality as measured by the California Psychological Inventory (CPI). The CPI was specifically developed for use in the workplace. Instead of focusing on a few broad domains, the CPI is designed to provide specific information that can be organized around the more general domains and to incorporate information from other assessments (Cobb et al., 2020). As such, the CPI focuses on facets of personality that provide nuance to other, more general, personality constructs such as the Big 5. The CPI was also designed to be easier to interpret and apply to workplace interventions and has the added advantage of being able to be completed in roughly 25 to 35 minutes. The characteristics measured by the CPI can provide insight into specific facets of personality as they relate to self-other agreement. Examining the linkages between the personality facets measured by the CPI and self-other agreement has the potential to provide two significant contributions to the use of 360-degree feedback in organizations. First, the crucial role of ratings in organizational decisions necessitates that we

understand what influences them. Second, organizations have long used personality measures as selection tools and understanding how those personality measures can potentially predict organizational outcomes is of significant importance.

The current study focuses on investigating and extending research examining individual personality characteristics related to rating tendencies. In this vein, factors influencing agreement between self and other ratings were analyzed using a self-report instrument measuring facets of work-related personality. Specifically, the current study focused on narrow aspects of personality as opposed to broad personality traits in answer to Barrick et al.'s (2001) call to investigate specific levels of analysis between performance appraisal and personality. Focusing on narrow traits will enable interventions directed at specific personality influences in performance feedback and leadership development. Specific personality variables linked to rating tendencies, followed by the hypotheses proposed to be tested, are described next. Since there is a dearth of research on the CPI variables' relationships with self-other agreement, these hypotheses were developed based on the extant literature exploring other measures of personality where evidence of theoretical and empirical links existed between CPI personality facets and other personality variables, such as the Big 5, that have been shown to have links with rating congruence.

### *Dominance and Assertiveness*

Dominance, as measured by the CPI, is characterized as “prosocial interpersonal dominance, strength of will, and perseverance in pursuit of goals” (Gough & Bradley, 2005, p. 6). Brutus et al. (1999), in an exploratory study examining linkages between personality and rating tendencies using the CPI, found that dominance differentially predicted rating agreement between participants’ ratings and ratings given by others. The authors posited that dominance influences how managers evaluate their leadership and behavior, impacting their self-ratings.

Dominance has also been found to correlate strongly with the Big 5 trait of Extraversion ( $r = .55$ ; McCrae et al., 1993). Judge et al. (2006), using the 44-item Big Five Inventory, reported that individuals high in extraversion were more likely to attach higher ratings to their performance. Visser et al. (2008) found similar results in a study where participants were asked to rate their level of intelligence. Two other studies, based on assessment center task performance, found extraverts, as measured by Goldberg’s 100 unipolar markers, and participants scoring high in dominance, as measured by the 15FQ+ personality questionnaire, were more likely to rate themselves higher than assessor ratings of their performance (Bell & Arthur Jr., 2008; Jackson et al., 2007).

Dominance has links to other essential personality facets measured by well-established personality instruments. The Oregon Research Institute (2022) reports



that dominance is linked to the International Personality Item Pool (IPIP) scale of assertiveness ( $r = .92$ ). Self-Verification theory, under the Social Relations Model, may provide support for assertiveness' role in rating consensus. Self-verification proposes that individuals aim to compel others to see them as they see themselves and that the stronger an individual feels about their self-concept, the less likely they are to be open to feedback contrary to their views (Kenny, 1994; Swann et al., 1989). From a theoretical perspective, individuals who are high in self-verification would then be likely to be low in self-other agreement.

Hypothesis 1: Individuals high in Dominance will overrate their performance when compared to supervisor, peer, and direct report ratings.

### *Sociability*

Individuals who are high in Sociability, as measured by the CPI, are “outgoing and socially affiliative, and who enjoy social participation” (Gough & Bradley, 2005, p. 6). Various measures of Sociability have been found to play a role in self-perception (Yammarino & Atwater, 1993). Wymer and Penner (1985) found that certain aspects of Sociability and social skills were related to rating congruence in a study using 228 undergraduate psychology students. Participants who scored higher in social skills and other-directedness were found to over-rate themselves relative to peers.

Sociability has been studied through various instruments, and it has been generally accepted that someone high in sociability is drawn to other people, enjoys meeting new people, and is socially competent (Boyle et al., 2008). The theoretical link between sociability and self-other agreement likely occurs through an interaction between extraversion, self-esteem, and social competence leading actors to attempt to participate in self-verification or self-enhancement. Weinstein (1969) defined social competence as the ability to manipulate others' responses. Sociability has been found to strongly correlate with the Big 5 factor of extraversion ( $r = .64$ ; McCrae et al., 1993). Extraversion has been found to have moderate correlations with self-esteem ( $r = .40$ ; Robins et al., 2001). Vaughan-Johnston et al. (2021) posited that extraversion was related to self-esteem when individuals engaged in self-enhancement strategies. As mentioned, the self-enhancement theory posits that individuals focus more on their positive and downplay their negative aspects (Kenny, 1994). Self-verification theory posits that individuals will attempt to convince others to see them as they see themselves. Thus, I propose that someone high in sociability would likely have higher extraversion and self-esteem and would therefore be more likely to rate themselves higher than others would rate them.

Hypothesis 2: Individuals high in Sociability will overrate their performance when compared to supervisor, peer, and direct report ratings.

### *Well Being*

As measured by the CPI, well-being refers to “feelings of physical and psychological well-being” (Gough & Bradley, 2005, p. 6). Well-being has also been defined as an individual’s evaluation of their lives relative to their “optimal psychological functioning and experience” (Ryan & Deci, 2001, p. 142). It should be noted that well-being, as measured by the CPI, refers to the definition of well-being as the “expectation to attain (and ultimately attaining) the outcomes one values” (Ryan & Deci, 2001, p. 145). This view of well-being is considered a more stable, longer-term trait-based view as opposed to the state-based construct measured by the subjective well-being (SWB) instrument most commonly associated with the term (Ryan & Deci, 2001; Ryff & Singer, 1998). It should also be noted that the use of SWB to measure well-being is controversial in some social psychological circles precisely because of its subjective focus (Ryff & Singer, 1998).

In a review of relationships between the CPI and the Big 5 personality traits, McCrae et al. (1993) found a moderate negative relationship between well-being and neuroticism ( $r = -.45$ ). Furnham et al. (2005) found that neuroticism was negatively related to self-estimates of intelligence but unrelated to intelligence scores. Well-being has also been found to be strongly correlated with the scale optimism in the IPIP (Oregon Research Institute, 2022). To date, only one study has directly examined well-being in relation to self-other agreement. In an

exploratory analysis of the relationship, Brutus et al. (1999) found that well-being was significantly linked to subordinate performance ratings.

In a study examining the relationship between personality and mental health in minorities, Pyant and Yanico (1991) found that Well Being, as measured by the CPI, was negatively and strongly related to Depression ( $r = -.55$ ), as measured by the Beck Depression Inventory. Depression has been linked to rating agreement. Abramson and colleagues have shown that lightly or moderately depressed individuals are more likely to rate themselves similarly to other raters when compared to non-depressed individuals (Abramson & Andrews, 1982; Alloy et al., 1981; Alloy & Abramson, 1982; Ingram, 1989). The authors in these studies posit that breakdowns in self-enhancing processes led to a tendency to be more realistic when evaluating performance behavior. Depressed individuals may also avoid over-rating themselves to reduce criticism, to seek praise by offering lower ratings, or because they are more in tune with others (Atwater & Yammarino, 1997; Fleenor et al., 2010). In an early study of the CPI and agreement in ratings of intelligence, Hjelle (1969) found that individuals who scored higher in well-being were more likely to provide congruent ratings with their college roommates on ratings of personality traits. The link between well-being and performance ratings has not been examined. Considering well-being's relationships with Neuroticism and Depression and the evidence of links with rating congruence on individual traits, further investigation is warranted.

Hypothesis 3: Individuals high in Well Being will offer congruent performance ratings when compared to supervisor, peer, and direct report ratings.

### *Empathy*

Empathy, as measured by the CPI, is “understanding how others feel and think” (Gough & Bradley, 2005, p. 6). Individuals high on empathy are thought to have “superior interpersonal skills and good intellectual ability” (Gough & Bradley, 2005, p. 42). In an exploratory analysis of the relationship between personality and self-other agreement, Brutus et al. (1999) found that empathy was significantly linked to performance ratings and rating congruence. In fact, in their study, empathy was the only trait they examined that was a significant predictor of ratings across all sources. The authors posited that individuals high in empathy perceive themselves as better leaders and are perceived similarly by others. Since the CPI was developed for work contexts and that the empathy facet focuses on measuring how well an individual understands the feelings and thoughts of others, it is possible that the term “understand” extends to how others think and feel about an individual’s performance. The CPI’s inclusion of interpersonal skills and intelligence into the construct of empathy, along with the links to effective leadership as viewed by the self and others, lead to the likelihood that individuals high in empathy will likely lead to positive, congruent ratings of performance.

Hypothesis 4: Individuals high in Empathy will offer congruent performance ratings when compared to supervisor, peer, and direct report ratings.

#### *Good Impression*

Good impression, as measured by the CPI, refers to the tendency to exhibit “overly strong attempts to create a favorable impression” with high scorers, whereas low scorers will be identified as people whose “style of self-presentation emphasizes ingratiation and compliance” (Gough & Bradley, 2005, p. 6). Additionally, Gough and Bradley (2005) claim that good impression, as measured by the CPI, is intended to “identify attempts to give an overly favorable impression, to the extent of faking” (p. 43). Atwater and Yammarino (1997) posited that cognitive processes such as beliefs can impact self-perception and self-ratings. Self-ratings are also impacted by an individual’s interpretation of how others perceive them (Ashford, 1989). It is likely that individuals high in good impression engage in self-enhancement, self-presentation, and/or self-verification processes as their respective theories propose. Individuals engaging in self-enhancement focus more on their positive characteristics and downplay or ignore their negative ones (Kenny, 1994). Self-presentation refers to a tendency to view themselves in relation to how they want others to view them (Kenny, 1994). Finally, self-verification is when individuals attempt to convince others to see them as they see themselves (Kenny, 1994). Burusic and Ribar (2014) found that individuals are able to detect others’ self-presentation tactics, particularly self-promotion, with very little information.

Thus, it follows that individuals high in good impression will likely over-rate their performance compared to others.

Brutus et al. (1999) provided some evidence of a link between over-rating and good impression. In an exploratory study, they found differential prediction of ratings between the self, supervisor, and subordinates, wherein each rating group provided significantly lower ratings than the participant's. No other study has examined good impression as it relates to self-other agreement.

Hypothesis 5: Individuals high in Good Impression will overrate their performance when compared to supervisor, peer, and direct report ratings.

As stated previously, the above hypotheses were developed based on the extant literature on personality as it relates to performance appraisal. However, it is possible that links exist between other facets of personality, as measured by the CPI, and self-other agreement. Thus, rating tendencies as they relate to the remaining 21 facets of personality measured by the CPI were analyzed. No hypotheses were developed for these facets as the analysis is exploratory in nature.

## Chapter 3

### Method

#### Participants

The sample for this research consisted of 294 employees whose assessment scores and performance data were captured in an archival database by a management consulting firm. This consulting firm specializes in offering client solutions in the areas of selection, coaching, organizational development, and leadership development in a wide array of industries, including healthcare, construction, retail, hospitality, manufacturing, and government. The participants for the current study were identified by their organizations for leadership development opportunities such as coaching and training. No demographic data were available for the participants as the data were not collected in a manner that would facilitate including them in the analysis. The database included participant scores on personality and performance data captured using an online, custom-built 360-degree feedback system with ratings from four rating groups: self, supervisor, peer, and direct reports.



## Measures

### *California Personality Inventory (CPI)*

The personality assessment used in the current study is an online instrument known as the California Psychological Inventory. The CPI-260 consists of 260 true/false items and is offered in timed and untimed formats. Developed as a measure of common personality factors, the CPI is an industry-standard instrument used globally in business for selection and development purposes.

The CPI measures facets of personality across five primary dimensions using 26 scales (Gough & Bradley, 2005). The dimensions measured by the CPI include Interpersonal Orientation/Dealing with Others, Self-Management, Motivations and Cognitive Style, Personal Characteristics, and Work-Related Measures. The Interpersonal Orientation dimension describes how the individual relates and interacts with the interpersonal and social environment. The scales in this dimension include Dominance, Capacity for Status, Sociability, Social Presence, Self-Acceptance, Independence, and Empathy (Gough, 2009). The Self-Management dimension describes how the individual relates to the values and normative systems and their functioning within the normative frame of reference. The scales in this dimension include Responsibility, Socialization, Self-Control, Good Impression, Communality, Well-Being, and Tolerance. The Motivations and Cognitive Style dimension describes the individual's interest and orientation

toward practical or intellectual areas along with their personal preference of style of achievement. The scales in this dimension include Achievement via Conformism, Achievement Via Independence, and Conceptual Fluency. The Personal Characteristics dimension describes an individual's adherence to traditional role descriptions and how they analyze others. The scales in this dimension include Insightfulness, Flexibility, and Sensitivity. The Work-Related measures dimension describes the behavioral tendencies and preferences regarding the work environment and their working style. The scales in this dimension include Managerial Potential, Work Orientation, Creative Temperament, Leadership, Amicability, and Law Enforcement Orientation (Gough, 2009). Please see Table 1 for scale descriptions.

The CPI was developed initially in the early 1950s to assess enduring human personality traits (Cobb et al., 2020). It originally consisted of 434 self-report true-false items based on common terms used to describe personality as it relates to typical behavior patterns, opinions and feelings, and attitudes centered on familial, social, and ethnic concerns (Cobb et al., 2020). The CPI has been revised several times, with the latest version containing 260 items designed specifically for online administration in a work setting. The correlations between the CPI 434 and the CPI 260 range from .81 to .97. CPI reliability (alpha) median estimates for personality survey scores range from .70 to .76 with an overall median reliability of .66. The individual scales range from .43 for the Law Enforcement Orientation scale to .86

for the Dominance scale as calculated from 6,000 randomly selected cases (Gough & Bradley, 2005). The CPI reliability scores are comparable with those found on similar personality scales (Cobb et al., 2020; Gough & Bradley, 2005; Viswesvaran & Ones, 1999).

The scores of the CPI are normed and reported as T-scores from the publisher. T-scores represent a category of standardized scores computed by considering the scores obtained by other individuals from the normative sample. The mean of a T-score distribution is set at 50, having a standard deviation of 10 scores. The score  $T = 50$  indicates that the respondent can be described as similar to most individuals in the general population. The interval  $T = 40 - 60$  contains approximately 68% of the individual scores obtained in the normative sample. Theoretically, T scores range between  $T = 20$  and  $T = 80$ . The profile presents five interpretative categories for each scale. The interval  $T = 20 - 35$  is considered very low. The scores that fall in the interval  $T = 35 - 45$  are considered low (below average). The average interval ranges from  $T = 45$  to  $T = 55$ . The high category begins with  $T = 55$  and ends with  $T = 65$ . Finally, scores that fall in the interval  $T = 65 - 80$  are considered very high.

The CPI is of particular interest in this study for three reasons. First, the CPI 260 is based on common terminology, is easy to understand, and was developed specifically to be used in organizational psychology settings, resulting in it being one of the most commonly used personality measures in business (Camara et al.,

2000). Second, decades of research have supported the reliability and validity of the CPI and have provided an extensive examination of the CPI in relation to the Big 5 and other measures of personality, which will enable further understanding of the current research question in relation to other theories of personality and SOA. Third, an extensive search of the extant literature using Google Scholar, PsychInfo, and a review of the reference sections of relevant literature reviews and meta-analyses has revealed that, while there are over 60 studies examining the relationship between personality and SOA, no examination has been conducted using the CPI as the personality measure. Given its widespread use in business for employee selection and development, it is prudent to add this examination to the current body of knowledge surrounding the relationship between work personality characteristics and SOA.

### *Customized multi-rater instrument*

A team of Industrial and Organizational Psychologists developed the 360-degree feedback instrument used in the current study for commercial use. The instrument was designed to measure job performance across five dimensions: Analysis & Decision Making, Responsibility and Achievement, Relationships and Teaming, Managing and Developing, and Leading and Influencing. The five dimensions are measured by 80 items. Each dimension is measured by eight items focused on effective behaviors and eight items focused on potential derailers. Example items

for effective behaviors include “Demonstrates initiative and self-motivation,” “Monitors others' performance,” and “Delegates work effectively.” Examples of potential derailers include “Fails to hold others accountable,” “Takes credit for others' accomplishments,” and “Submits inaccurate or incomplete work.” Items are scored on a 9-point Likert scale. See Table 2 for 360-degree instrument scale descriptions.

The consulting management firm developed the instrument by reviewing the extant leadership literature to identify high and low-performing behaviors. The survey results were refined into five themes and 80 items through a Q-sort process. The final version was then piloted using 143 raters. The pilot study results were found to have adequate reliability, with alphas ranging from .77 to .90 (Colarelli, Meyer & Associates, Inc., 2012; Cortina, 1993).

Participants in the pilot study also completed the Multifactor Leadership Questionnaire (MLQ) regarding the ratee's transformational leadership behaviors. The scores of the 360-degree feedback measurement were correlated with the MLQ results. Correlations between the four dimensions of transformational leadership and the five dimensions of the 360-degree feedback instrument ranged between .31 and .79 and averaged at .56 for effective behaviors (Colarelli, Meyer & Associates, Inc., 2012). Correlations between the four dimensions of transformational leadership and the five dimensions of the 360-degree feedback instrument ranged

between -.21 and .004 and averaged at -.13 for potential derailers (Colarelli, Meyer & Associates, Inc., 2012).

Finally, the pilot study participants were asked to rate the ratee on scales of overall effectiveness, promotability, and organizational agility. An ANOVA with post hoc analysis was conducted using those scores and scale scores of the effective behaviors and potential derailers. Overall, effective behaviors and potential derailers predicted 20% of each participant's overall effectiveness, 50% of promotability, and 52% of organizational agility ratings in the sample (Colarelli, Meyer & Associates, Inc., 2012).

### *Classification of Under/Over Rater*

The final measurement for the current study was statistically constructed to identify rating trends among self-raters. Based on Atwater & Yammarino's (1997) posited categories of self-raters as being under-raters, over-raters, or in-agreement raters, categorical variables were created using standardized difference scores. A point-biserial correlational analysis of these variables with personality facets was expected to provide insight into the likelihood of relative scores on personality facets falling into one of the rating tendency categories.

## Procedures

Both performance and personality data were collected online. The human resources departments or organizational leadership of each of the consulting management firm's clients identified the participants for development for promotion, succession planning, team building, leadership, and personal development. The participants completed the CPI as part of a series of five assessments. The participants were then asked to rate their own performance for the 360-degree feedback process. The participants' human resources departments and supervisors were asked to identify appropriate raters among the participant's supervisors, peers, and direct reports to complete ratings on the participants' performance. The data were collected to link the participants' personality and other characteristics to the five work-related competencies to facilitate the crafting of individual development plans. All participants and raters were informed of the purpose of the data collection before completing the instruments. All individuals were trained on the 360-degree feedback instrument before completing it and instructed to be as candid as possible. All responses and personal identifying information were kept confidential and secure by the third-party management consultant firm. Ratings were reported to participants only if three or more responses were received.

## Chapter 4

### Results

#### Preliminary Analyses

Scale scores were created indicating mean scores across all five dimensions of performance combined for each of the four rating groups: self ( $n = 294$ ), supervisor ( $n = 414$ ), peers ( $n = 1,160$ ), and direct reports ( $n = 1,037$ ). A separate scale score was also calculated for all “other” raters that indicated the mean of the ratings across the supervisor, peer, and direct rating groups. The scale scores encapsulated overall performance. Additionally, dummy coded variables were created to indicate rating tendencies by calculating rating tendencies using difference scores.

Difference scores were calculated by subtracting each rating group’s scores from participants’ self-rating scores. “Overraters” and “underraters” were identified by those whose difference scores fell outside of one-half standard deviation from zero (Fleenor et al., 2010; Sinha et al., 2012). All difference scores within one-half standard deviation from zero indicated “in agreement.”

Preliminary analysis of the collected data included examining the means, standard deviations, and correlations among the personality domain, performance, and rating tendency variables. See Table 3. Overall performance rating means tended to be



somewhat high with a low variance. However, this is expected given that the sample consisted of identified high-potential employees and the evidence that suggests higher ratings when the supervisor ratings are not anonymous (Atwater et al., 1998). Correlations between self-ratings and supervisor ( $r = .12, p < .05$ ), peer ( $r = .16, p < .05$ ), direct report ( $r = .18, p < .01$ ), and combined other ( $r = .17, p < .01$ ) ratings were low, which is consistent with prior research (Judge et al., 2006; Mabe & West, 1982; Ostroff et al., 2004).

## Hypothesis Testing

Historically, self-other agreement has been analyzed by calculating a difference score and analyzing the variance associated with other variables or regressing the difference score on potential predictors. However, this method has been criticized for its inability to be interpreted unambiguously (Edwards, 1995). Difference scores cannot elucidate how an independent variable impacts the self or the other rating individually. Thus, Edwards (1995) suggested, when self-other agreement was examined as a dependent variable, that multivariate regression was a more suitable analysis as it allows both the self-score and the other-score to be regressed simultaneously onto the independent variables. This approach enables the analysis of the predictor's impact on each rating while taking the other rating into account. For example, if self and supervisor ratings were regressed onto a personality trait, such as neuroticism, multivariate regression allows an interpretation of the overall

model to determine whether a relationship exists between the ratings and neuroticism. This omnibus multivariate test is based on Wilks'  $\Lambda$  and tests the two equations (i.e., self-ratings and supervisor ratings) jointly. A significant F statistic in the omnibus analysis indicates that the relationships between personality predictors and each dependent variable (self, supervisor, etc.) can be examined to assess the nature of the agreement effects. A non-significant result on the omnibus indicates that no relationship exists between the personality trait and either of the performance ratings. The analysis then presents regression coefficients for each rating for individual interpretation, therefore facilitating inferences about how the independent variable affects each rating when the other rating is considered. Specifically, if the regression coefficient for the self-rating variable is significant, then it can be inferred that the score of the personality variable has a significant impact on self-ratings. Similar inferences can be made for the supervisor, peer, direct report, and combined "other" rating variables. Thus, if both equations are significant, then the personality scores can be assumed to have an impact on both sets of ratings in the regression. In this case, an examination of the coefficients will lead to insights as to how the personality variable affects self-other rating agreement. If only one equation is statistically significant, then the direction (positive/negative) and strength of the coefficient is examined to provide evidence of over-rating (positive self-rating or negative other-rating) or under-rating (negative self-rating or positive other-rating). Any finding of significance on one

variable and not the other will likely suggest over- or under-rating at some point in the personality variable spectrum.

To examine the hypotheses in the current study, twenty multivariate regressions were conducted using the four rating groups' ratings of overall performance as the dependent variables and the personality variables as the independent variable. Specifically, the dependent variables in each regression consisted of the self-rating scale score paired with each rating group scale score (self-supervisor, self-peer, self-direct reports, self-mean of all "other" raters). The independent variable was one of the five personality dimensions of the CPI identified in the hypotheses. Each multivariate regression examined the relationship between rating tendency and personality factors. The results of these multivariate tests are contained in Table 4. These multivariate regressions were supplemented by also examining the relationships between the CPI dimensions and the classification of being an over or under rater.

Hypothesis 1 predicted that participants who scored higher on dominance would overrate their performance compared to supervisor, peer, and direct report ratings. Point biserial correlations indicated that the higher a participant scored on dominance, the less likely they were to be identified as having under-rating tendencies ( $r = -.14, p < .05$ ) while having no significant relationship with overrating ( $r = .09, p > .05$ ) or in-agreement rating ( $r = .06, p > .05$ ) tendencies

when compared to supervisor ratings. Participants scoring higher on dominance were more likely to over-rate ( $r = .13, p < .01$ ) and less likely to under-rate ( $r = -.24, p < .01$ ) while having no significant relationship with in-agreement ratings ( $r = .06, p > .05$ ) compared to peer ratings. Participants scoring higher on dominance were more likely to over-rate ( $r = .21, p < .01$ ) or provide in-agreement ratings ( $r = .15, p < .01$ ) and less likely to under-rate ( $r = -.22, p < .01$ ) compared to direct report ratings. Finally, participants scoring higher on dominance were more likely to over-rate ( $r = .20, p < .01$ ) and less likely to under-rate ( $r = -.22, p < .01$ ) while having no significant relationship with in agreement ( $r = .09, p > .05$ ) rating tendencies in relation to combined other ratings.

Results of the multivariate regressions indicated that dominance was significantly related to the sets self- and supervisor-ratings (Wilks'  $\Lambda = .89, F(2, 275) = 17.22, p < .001$ ), self- and peer-ratings (Wilks'  $\Lambda = .89, F(2, 238) = 14.44, p < .001$ ), self- and direct report-ratings (Wilks'  $\Lambda = .87, F(2, 215) = 16.01, p < .001$ ), and self- and combined other-ratings (Wilks'  $\Lambda = .89, F(2, 291) = 17.44, p < .001$ ). The canonical correlations indicated that dominance accounted for 11.1% of the variance in the set of self-supervisor ratings, 10.8% of the variance in the set of self-peer ratings, 13.0% of the variance in the set of self-direct report ratings, and 10.7% of the variance in the set of self-combined other ratings.

Four multivariate regressions were conducted using self- ( $\beta = .033, p < .001$ ) and supervisor ratings ( $\beta = .011, p > .05$ ), self- ( $\beta = .032, p < .001$ ) and peer ratings ( $\beta = -.002, p > .05$ ), self- ( $\beta = .035, p < .001$ ) and direct report ratings ( $\beta = -.001, p > .05$ ), and self- ( $\beta = .033, p < .001$ ) and combined other ratings ( $\beta = .003, p > .05$ ) as the dependent variables that were simultaneously regressed on the dominance scores in the sample. The regression coefficients indicated that dominance was significantly positively related to self-ratings and was not significantly related to supervisor, peer, direct report, and combined other ratings. Since dominance had no significant relationship in the regression analyses with supervisor, peer, direct report, or combined ratings, there is no evidence that dominance differentially influences “other” ratings in this sample. This indicates that individuals higher in dominance will provide higher ratings than all other rating groups. The point biserial correlation results above indicated that dominance had a significant relationship with over-rating tendencies compared to peers, direct reports, and combined other ratings. Thus, hypothesis 1 was supported.

Hypothesis 2 predicted that participants who scored higher on sociability would overrate their performance compared to supervisor, peer, and direct report ratings. Point biserial correlations indicated that the higher a participant scored on sociability, the less likely they were to be identified as having under-rating tendencies ( $r = -.16, p < .05$ ) while having no significant relationship with overrating ( $r = .10, p > .05$ ) or in-agreement rating ( $r = .05, p > .05$ ) tendencies

when compared to supervisor ratings. Participants scoring higher on sociability were less likely to under-rate ( $r = -.17, p < .05$ ) while having no significant relationship with overrating ( $r = .05, p > .05$ ) or in-agreement ratings ( $r = .02, p > .05$ ) compared to peer ratings. Participants scoring higher on sociability were less likely to under-rate ( $r = -.22, p < .01$ ) and more likely to present overrating ( $r = .13, p < .01$ ) or in-agreement rating ( $r = .19, p < .01$ ) tendencies compared to direct report ratings. Finally, participants scoring higher on sociability were less likely to under-rate ( $r = -.22, p < .01$ ) and more likely to over-rate ( $r = .16, p < .05$ ) or provide in-agreement ratings ( $r = .12, p < .05$ ) compared to combined other ratings.

Results of the multivariate regressions indicated that sociability was significantly related to the sets self- and supervisor-ratings (Wilks'  $\Lambda = .92, F(2, 275) = 12.73, p < .001$ ), self- and peer-ratings (Wilks'  $\Lambda = .92, F(2, 238) = 10.23, p < .001$ ), self- and direct report-ratings (Wilks'  $\Lambda = .86, F(2, 215) = 14.01, p < .001$ ), and self- and combined other-ratings (Wilks'  $\Lambda = .92, F(2, 291) = 13.54, p < .001$ ). The canonical correlations indicated that sociability accounted for 8.5% of the variance in the set of self-supervisor ratings, 8.0% of the variance in the set of self-peer ratings, 11.5% of the variance in the set of self-direct report ratings, and 8.5% of the variance in the set of self-combined other ratings.

Four multivariate regressions were conducted using self- ( $\beta = .025, p < .001$ ) and supervisor ratings ( $\beta = .004, p > .05$ ), self- ( $\beta = .024, p < .001$ ) and peer ratings ( $\beta =$

.005,  $p > .05$ ), self- ( $\beta = .030, p < .001$ ) and direct report ratings ( $\beta = .002, p > .05$ ), and self- ( $\beta = .025, p < .001$ ) and combined other ratings ( $\beta = .004, p > .05$ ) as the dependent variables that were simultaneously regressed on the sociability scores in the sample. The regression coefficients indicated that sociability was significantly positively related to self-ratings and was not significantly related to supervisor, peer, direct report ratings, and combined other ratings. Since sociability had no significant relationship in the regression analyses with supervisor, peer, direct report, or combined ratings, there is no evidence that sociability differentially influences “other” ratings in this sample. This indicates that individuals higher in sociability will provide higher ratings than all other rating groups. The point biserial correlation results above indicated that sociability had a significant relationship with over-rating tendencies compared to direct reports and combined other ratings. Thus, hypothesis 2 was supported.

Hypothesis 3 predicted that participants who scored higher on well-being would offer congruent performance ratings compared to supervisor, peer, and direct report ratings. Point biserial correlations indicated that participants scoring higher on well-being were less likely to present under-rating tendencies ( $r = -.12, p < .05$ ) while having no significant relationship with over-rating ( $r = .09, p > .05$ ) or in-agreement rating ( $r = .01, p > .05$ ) tendencies when compared to supervisor ratings. Participant scores on well-being had no significant relationship with over-rating ( $r = .05, p > .05$ ), under-rating ( $r = -.06, p > .05$ ), or in-agreement ratings ( $r = .00, p >$

.05) compared to peer ratings. Participants scoring higher on well-being were more likely to over-rate ( $r = .12, p < .05$ ) or provide in-agreement ratings ( $r = .12, p < .05$ ) while having no significant relationship with under-rating ( $r = -.09, p > .05$ ) tendencies compared to direct report ratings. Finally, participants scoring higher on well-being were less likely to under-rate ( $r = -.18, p < .01$ ) or over-rate ( $r = .12, p < .05$ ) while having no significant relationship with in-agreement rating tendencies ( $r = .11, p > .05$ ) compared to combined other ratings.

Results of the multivariate regressions indicated that well-being was significantly related to the sets of self- and supervisor-ratings (Wilks'  $\Lambda = .90, F(2, 275) = 15.35, p < .001$ ), self- and peer-ratings (Wilks'  $\Lambda = .90, F(2, 238) = 12.75, p < .001$ ), self- and direct report-ratings (Wilks'  $\Lambda = .93, F(2, 215) = 7.46, p < .001$ ), and self- and combined other-ratings (Wilks'  $\Lambda = .90, F(2, 291) = 16.16, p < .001$ ). The canonical correlations indicated that well-being accounted for 10.0% of the variance in the set of self-supervisor ratings, 9.7% of the variance in the set of self-peer ratings, 6.5% of the variance in the set of self-direct report ratings, and 10.0% of the variance in the set of self-combined other ratings.

Four multivariate regressions were conducted using self- ( $\beta = .039, p < .001$ ) and supervisor ratings ( $\beta = .014, p > .05$ ), self- ( $\beta = .038, p < .001$ ), and peer ratings ( $\beta = .014, p > .05$ ), self- ( $\beta = .035, p < .001$ ) and direct report ratings ( $\beta = .002, p > .05$ ), and self- ( $\beta = .040, p < .001$ ) and combined other ratings ( $\beta = .011, p > .05$ ) as



the dependent variables that were simultaneously regressed on the well-being scores in the sample. The regression coefficients indicated that well-being was significantly positively related to self and was not significantly related to supervisor, peer, direct report, and combined other ratings. Since well-being had no significant relationship in the regression analyses with supervisor, peer, direct report, or combined ratings, there is no evidence that sociability differentially influences “other” ratings in this sample. However, the point biserial correlation results above indicated that well-being had a significant relationship with in-agreement rating tendencies compared to direct reports ratings. Thus hypothesis 3 was partially supported.

Hypothesis 4 predicted that participants who scored higher on empathy would offer congruent performance ratings compared to supervisor, peer, and direct report ratings. Point biserial correlations indicated no significant relationship existed between empathy and over-rating, under-rating, or in-agreement rating tendencies when compared to supervisor, peer, direct report, or combined other ratings.

Results of the multivariate regressions indicated that empathy was significantly related to self- and combined other-ratings (Wilks'  $\Lambda = .98$ ,  $F(2, 291) = 3.26$ ,  $p < .05$ ) and indicated no significant relationships with self- and supervisor-ratings, self- and peer-ratings, and self- and direct report-ratings. The canonical correlations

indicated that empathy accounted for 2.2% of the variance in the set of self-combined other ratings.

The regression coefficients indicated that empathy was significantly positively related to self-ratings ( $\beta = .012, p < .001$ ) when regressed with combined other ratings ( $\beta = .000, p > .05$ ). Empathy did not differentially predict between self-ratings and any other rating group. The point biserial correlation results above indicated that empathy had no significant relationships with over-, under-, or in-agreement rating tendencies in any rating group. Thus, hypothesis 4 was not supported.

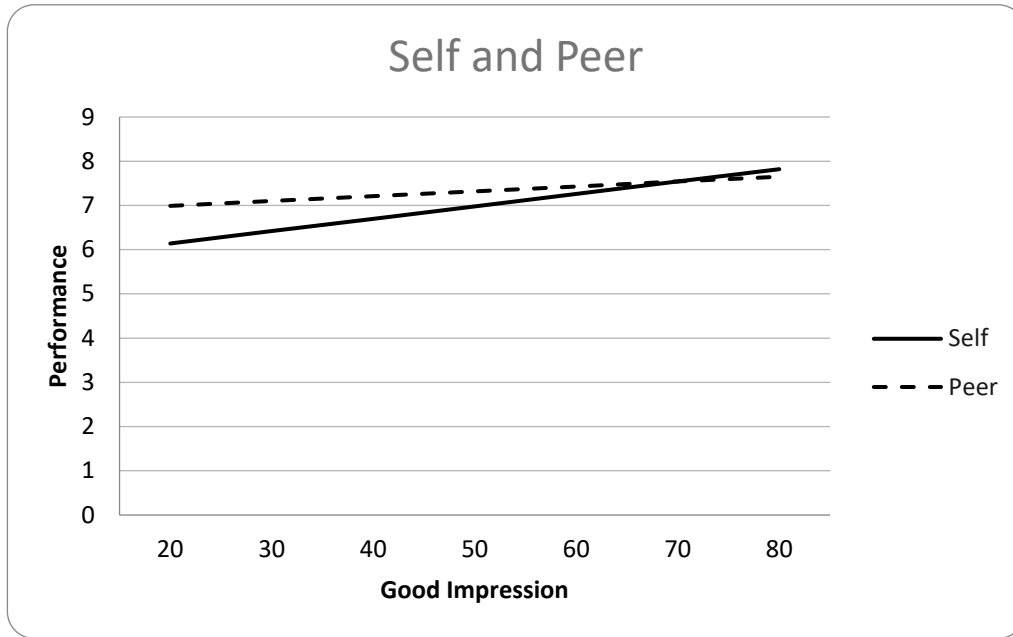
Hypothesis 5 predicted that participants who scored higher on good impression would overrate their performance compared to supervisor, peer, and direct report ratings. Point biserial correlations indicated that the higher a participant scored on good impression, the more likely they were to be identified as having over-rating tendencies ( $r = .16, p < .01$ ) and less likely to have under-rating ( $r = -.12, p > .05$ ) tendencies while having no significant relationship with in-agreement rating ( $r = .08, p > .05$ ) tendencies when compared to supervisor ratings. Participants scoring higher on good impression were more likely to over-rate ( $r = .14, p < .05$ ) and less likely to under-rate ( $r = -.12, p < .05$ ) while having no significant relationship with in-agreement ratings ( $r = -.02, p > .05$ ) compared to peer ratings. Participants scoring higher on good impression were more likely to over-rate ( $r = .23, p < .01$ )

and less likely to under-rate ( $r = -.15, p < .01$ ) while having no significant relationship with in-agreement rating ( $r = .06, p > .05$ ) tendencies compared to direct report ratings. Finally, participants scoring higher on good impression were more likely to over-rate ( $r = .17, p < .01$ ) and less likely to under-rate ( $r = -.15, p < .05$ ) while having no significant relationship with in-agreement ratings ( $r = .00, p > .05$ ) compared to combined other ratings.

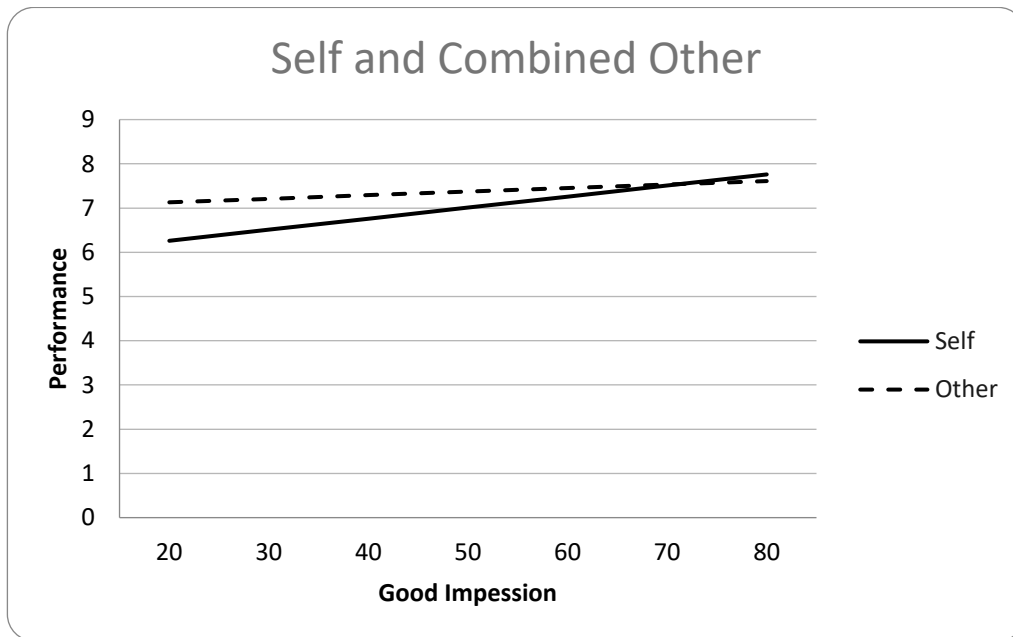
Results of the multivariate regressions indicated that good impression was significantly related to self- and supervisor-ratings (Wilks'  $\Lambda = .88, F(2, 275) = 18.42, p < .001$ ), self- and peer-ratings (Wilks'  $\Lambda = .84, F(2, 238) = 22.01, p < .001$ ), self- and direct report-ratings (Wilks'  $\Lambda = .89, F(2, 215) = 13.23, p < .001$ ), and self- and combined other-ratings (Wilks'  $\Lambda = .87, F(2, 291) = 21.71, p < .001$ ). The canonical correlations indicated that good impression accounted for 11.8% of the variance in the set of self-supervisor ratings, 15.6% of the variance in the set of self-peer ratings, 11.0% of the variance in the set of self-direct report ratings, and 13.0% of the variance in the set of self-combined other ratings.

Four multivariate regressions were conducted using self- ( $\beta = .024, p < .001$ ) and supervisor ratings ( $\beta = .007, p > .05$ ), self- ( $\beta = .028, p < .001$ ) and peer ratings ( $\beta = .011, p < .05$ ), self- ( $\beta = .025, p < .001$ ) and direct report ratings ( $\beta = .003, p > .05$ ), and self- ( $\beta = .025, p < .001$ ) and combined other ratings ( $\beta = .008, p < .05$ ) as the dependent variables that were simultaneously regressed on the good impression

scores in the sample. The regression coefficients indicated that good impression was significantly positively related to self, peer, and combined other ratings and was not significantly related to supervisor and direct report ratings. As shown in Figure 1, good impression differentially predicted self ( $\beta = .028, p < .001$ ) and peer ( $\beta = .011, p < .05$ ) ratings, indicating that, while good impression had a positive relationship with both sets of ratings, participants higher in good impression tended to rate themselves higher than did their peers. As shown in Figure 2, good impression differentially predicted self ( $\beta = .025, p < .001$ ) and combined other ( $\beta = .008, p < .001$ ) ratings, indicating that, while good impression had a positive relationship with both sets of ratings, participants higher in good impression tended to rate themselves higher than did all other raters combined. The point biserial correlation results listed above indicated that good impression had a significant relationship with over-rating tendencies compared to supervisor, peer, direct report, and combined other ratings. Thus, hypothesis 5 was supported.



**Figure 1: The Effect of Good Impression on Self and Peer Ratings**



**Figure 2: The Effect of Good Impression on Self and Other Ratings**

### *Exploratory Analysis*

An exploratory analysis of the links between the remaining 21 personality facets and the 12 rating tendency categories was conducted using point bi-serial correlations. See Table 1 for the personality facets and their definitions. The rating tendency categories consisted of under-, over-, and in-agreement rating tendencies for each rating group (supervisor, peer, subordinate, combined other). Correlations ranged from .24 (sensitivity, under-rate direct report) to -.26 (insight, under-rate combined other). See Table 5.

Achievement via independence and creative temperament were not significantly related to any of the rating tendency categories. Individuals who scored higher in capacity for status ( $r = .13, p < .05$ ), independence ( $r = .13, p < .05$ ), responsibility ( $r = .13, p < .05$ ), social conformity ( $r = .17, p < .01$ ), self-control ( $r = .18, p < .01$ ), tolerance ( $r = .15, p < .01$ ), achievement via conformance ( $r = .15, p < .01$ ), conceptual fluency ( $r = .15, p < .01$ ), managerial potential ( $r = .16, p < .01$ ), work orientation ( $r = .13, p < .05$ ), leadership ( $r = .19, p < .01$ ), amicability ( $r = .13, p < .05$ ), and law enforcement orientation ( $r = .20, p < .05$ ) were significantly more likely to over-rate their performance compared to direct report ratings. Individuals who scored higher in capacity for status ( $r = -.22, p < .01$ ), social presence ( $r = -.14, p < .05$ ), self-acceptance ( $r = -.17, p < .01$ ), independence ( $r = -.26, p < .01$ ), responsibility ( $r = -.13, p < .05$ ), social conformity ( $r = -.13, p < .05$ ), tolerance ( $r = -$

.20,  $p < .01$ ), conceptual fluency ( $r = -.16, p < .01$ ), insightfulness ( $r = -.21, p < .01$ ), managerial potential ( $r = -.24, p < .01$ ), work orientation ( $r = -.14, p < .05$ ), leadership ( $r = -.23, p < .01$ ), and amicability ( $r = -.12, p < .05$ ) were significantly less likely to under-rate their performance compared to combined other ratings. Of note is that independence was the only personality variable, aside from those in the hypotheses, that significantly related to agreement with peer ratings ( $r = .14, p < .05$ ). Also of note is that social conformity ( $r = .12, p < .05$ ) and work orientation ( $r = .12, p < .05$ ) were the only personality variables, aside from those in the hypotheses, that significantly related to over-rating compared to supervisor ratings. Insightfulness ( $r = .14, p < .05$ ) was also the only variable, aside from those in the hypotheses, that significantly related to in-agreement ratings compared to supervisor ratings.

Individuals who scored higher in capacity for status were significantly less likely to under-rate compared to supervisor ( $r = -.13, p < .05$ ), peer ( $r = -.19, p < .01$ ), direct report ( $r = -.15, p < .01$ ), and combined other ratings ( $r = -.22, p < .01$ ), were more likely to over-rate compared to direct report ( $r = .13, p < .05$ ) and combined other ratings ( $r = .15, p < .05$ ), and were more likely to provide in-agreement ratings with direct report ratings ( $r = .12, p < .05$ ). Individuals who scored higher in self-acceptance were significantly less likely to under-rate compared to supervisor ( $r = -.13, p < .05$ ), peer ( $r = -.17, p < .01$ ), direct report ( $r = -.16, p < .01$ ), and combined other ratings ( $r = -.17, p < .01$ ), were more likely to over-rate compared to

combined other ratings ( $r = .17, p < .01$ ), and were more likely to provide in-agreement ratings with direct report ( $r = .20, p < .01$ ) and combined other ratings ( $r = .16, p < .01$ ). Individuals who scored higher in independence were significantly less likely to under-rate compared to supervisor ( $r = -.12, p < .05$ ), peer ( $r = -.22, p < .01$ ), direct report ( $r = -.15, p < .01$ ), and combined other ratings ( $r = -.26, p < .01$ ), were more likely to over-rate compared to peer ( $r = .13, p < .05$ ) and subordinates ( $r = .14, p < .05$ ), and were more likely to provide in-agreement ratings with peers ( $r = .14, p < .05$ ) and combined others ( $r = .13, p < .05$ ).

Individuals who scored higher in sensitivity were significantly less likely to over-rate peer ( $r = -.15, p < .01$ ), subordinate ( $r = -.23, p < .01$ ), and combined all other ratings ( $r = -.21, p < .01$ ) and more likely to under-rate compared to supervisor ( $r = .13, p < .05$ ), direct reports ( $r = .24, p < .01$ ), and combined other ratings ( $r = .22, p < .01$ ). Individuals who scored higher in Managerial Potential were significantly less likely to under-rate compared to supervisor ( $r = -.19, p < .01$ ), peer ( $r = -.22, p < .01$ ), direct report ( $r = -.14, p < .05$ ), and combined other ratings ( $r = -.24, p < .01$ ), were more likely to provide ratings in-agreement with direct reports ( $r = .17, p < .01$ ) and combined other ratings ( $r = .14, p < .05$ ), and were more likely to over-rate compared to peer ( $r = .13, p < .05$ ) and combined other ratings ( $r = .13, p < .05$ ). Finally, individuals who scored higher in leadership were significantly less likely to under-rate compared to supervisor ( $r = -.18, p < .01$ ), peer ( $r = -.25, p < .01$ ), direct report ( $r = -.17, p < .01$ ), and combined all other ratings ( $r = -.23, p < .01$ ), were



more likely to provide in-agreement ratings with direct reports ( $r = .17, p < .01$ ) and combined other ratings ( $r = .12, p < .05$ ), and were more likely to over-rate compared to subordinate ( $r = .19, p < .01$ ) and combined other ratings ( $r = .21, p < .01$ ).

## Chapter 5

### Discussion

This study aimed to investigate the links between self-other rating tendencies and narrow facets of personality. Overall job performance ratings were examined to further our understanding of how personality traits influence them. Ratings across four different rating groups were assessed for evidence of differential predictive ability among certain personality facets among raters in different roles. In addition, an exploration of all personality dimensions measured by the CPI was conducted to identify potential links with rating tendencies.

Hypotheses 1 and 2 posited that individuals with higher scores in dominance and sociability would be related to over-rating their job performance compared to supervisors, peers, and direct report ratings. Initial analysis indicated that individuals higher in dominance were more likely to over-rate their performance compared to peers, direct reports, and combined other ratings. Individuals high in sociability were more likely to over-rate compared to subordinates and combined other scores. In a deeper examination of these hypotheses for differential prediction, the personality facets of dominance and sociability indicated positive relationships with self-ratings when examined in conjunction with ratings from all rating groups. These findings indicate that the higher one scores on dominance or

sociability, the higher their self-ratings of job performance will be. However, there were no associations with any other rating groups' ratings indicated, suggesting that, while dominance and sociability influence self-ratings across the board, there is no evidence of direct influence on the ratings of others. Overall, it is clear that dominance and sociability impact self-ratings to the point where over-rating is likely if for no other reason than a change in self-ratings along with no associated change in other ratings will eventually lead to divergence and less rating agreement. In addition, these findings offer no support for the assertion that individuals high in dominance and sociability participate in self-verification behaviors, as posited by Kenny (1994) and Swann et al. (1989), or if they do, there is no evidence of successfully changing others' views on their performance.

Hypothesis 3 proposed that individuals high in well-being would offer ratings similar to those of others. The results indicate similar findings as the previous two hypotheses. The initial analysis indicated a similarly higher likelihood of presenting congruent ratings or over ratings in comparison to subordinates when well-being is high. Additionally, those higher in well-being were less likely to be under-raters when compared to supervisor ratings, supporting previous findings (Hjelle, 1969; Pyant & Yanico, 1991). Well-being was also positively associated with performance ratings in comparison with all rating groups and had no association with those rating groups' scores. These findings partially support and extend the findings of

Brutus et al. (1999) in confirming that well-being can impact how individuals are viewed by their subordinates in ways that supervisors and peers do not share.

Hypothesis 4 posited that individuals higher in empathy would offer ratings that were more in agreement with supervisors, peers, and direct reports. Contrary to the findings of previous research (i.e., Brutus et al., 1999), no evidence of links between empathy and self-other agreement was found. One possible explanation of these results is that they are an artifact of the sample used in this study. Participants were identified as high-potential employees. Another possible explanation is that the CPI's definition of empathy may not extend to performance. Since the item level information was not made available from the licensing organization, there is currently no method for exploring the factors measured in the empathy domain. Additionally, no demographic information was available to analyze for covariation. Future studies should examine self-other agreement with a broader range of performance behavior and carefully analyze potential covariates and interactions to provide further clarity on any potential links between empathy and rating congruence.

Hypothesis 5 proposed that individuals high in good impression would over-rate their performance compared to supervisor, peer, and direct report ratings. The initial analysis indicated that those high in good impression were more likely to have over-rating tendencies when compared to supervisor, peer, direct report, and

combined other ratings. In a deeper analysis, good impression indicated positive relationships with self-ratings when examined in conjunction with ratings from all rating groups. Additionally, good impression was the only variable examined in this study that indicated a relationship with the ratings of other raters. Peer ratings and the combined ratings of all raters were shown to have a positive relationship with good impression. These relationships were weaker than for self-ratings. When examining the results of all analyses, higher good impression leads to not only higher self, peer, and combined other ratings but also to over-rating tendencies. These findings support Atwater and Yammarino's (1997) and Ashford's (1989) propositions that self-perception components, such as self-enhancement, self-presentation, and self-verification, influence self-other agreement.

## Theoretical Implications

With a few notable exceptions (see Sinha, 2012; Vries et al., 2016), examinations of the influence of personality on self-other rating agreement are conspicuously scarce in the literature. The current study builds on a handful of studies (see Fleenor et al., 2010, Brutus et al., 1999) connecting personality facets to self-other agreement research, exploring how an individual's personality affects self- and other-ratings of their performance behaviors. The current study advances the knowledge of rating congruence in the following ways. First, the current study highlighted the importance of applying more appropriate methods to study the

prediction of self-other agreement (e.g., Edwards, 1995) in that such methods are suitable for depicting a more refined understanding of self-other agreement that has not yet received enough attention. Whereas other studies have associated general personality factors with self-other agreement, the current study explored how narrow facets of personality can potentially impact rating agreement. Even though some personality facets do not predict other-ratings, this does not imply that they have no impact on how others view an individual's performance in light of their personality. Thus, theories of self-other agreement should be revisited to consider the complex multivariate links of how personality interacts with perceptions of behavior.

Second, the findings in the current study support the concept that various aspects of personality differentially affect both self- and other-ratings of performance, which produce varying degrees of self-other agreement. This provides a deeper understanding of how individual differences such as personality affect performance ratings from multiple perspectives, including rating source, self-enhancement, self-verification, and self-presentation. Current theory may need to be refined to consider these perspectives. Theory and research may also want to revisit how personality facets may predict self-other agreement in other ways that may not follow the typical main effects findings examined in isolation.

## Practical Implications

The findings of research into the antecedents of self-other rating agreement can be leveraged across a wide range of talent management applications. Self-other agreement, and factors that can predict it, are rarely considered in leader development and selection systems. Organizations have an opportunity to operationalize findings from research on the antecedents of self-other agreement to tailor development and training interventions along with selection systems. For example, under-raters have been found to have lower organizational and lower compensation levels and have been posited to be less likely to pursue challenging tasks or persists in those tasks for which they are qualified due to low or negative self-perceptions (Ostroff et al., 2004). The finding that a person who is low in dominance is more likely to under-rate their performance could be targeted for interventions designed to empower employees and adjust their self-perception to match their potential performance. A well-developed personality profile of individuals likely to provide in-agreement ratings can also assist organizations in identifying, hiring, and selecting high-potential employees for further development.

## Limitations

Due to the field study nature of the current research, there are several limitations. First, the study sample comprised individuals selected for leadership development by their organizations and was not random. This implies that the participants were

high-potential employees with a history of good job performance. This likely impacted the analysis by providing range-restricted job performance ratings, as there were no “poor” performers in the sample. The means ( $M = 7.46$ ) and standard deviations ( $SD = .489$ ) of the overall performance ratings indicate the potential for significant range restriction. It is also possible that the personality facets measured in the current study could be restricted as the characteristics of a high-potential employee may differ from other populations. The focus outcome variable of the current study was rating consensus, which might also be impacted as it is possible that the sample was selected for leadership development due to some variables not captured in the current data set, including a tendency to provide accurate self-ratings that are in agreement with other ratings. As such, it is important to assess the current study's findings as indicative of the relationships within the limited range of the scores obtained and has little predictive power for values outside that range.

Second, the current study is cross-sectional in nature and does not provide clear insight into how personality facets can impact performance, and the ratings thereof, over time. Parallel to this, no tenure data were collected from the current sample. Research has provided some evidence that tenure plays a role in self-other agreement (Fleenor et al., 2010). Thus, the current study's findings should be viewed as a snapshot in time and, therefore, not predictive of long-term links between personality and self-other agreement.



Finally, several limitations in the available information existed. For example, personality and performance data were collected for the participant only. No further information was collected regarding the participants, such as demographic data, tenure, or organizational level. Also, no data were collected regarding the demographic and personality traits of the raters in any of the other rating groups, which may also play a role in rating agreement (Atwater & Yammarino, 1997; Brutus et al., 1999; Fleenor et al., 2010). This information may have been useful to determine potential covariates of the links explored here and present a potential “third variable” problem.

## Future Research Directions

Future studies examining self-other agreement as it relates to performance and personality should attempt to gather as much information as possible about the participants and the raters. In addition, ratings from other sources, such as customers and other potential rating sources, should be examined to provide further insight. For example, exploring ratings of individuals who work in a remote vs. in-person context as the “observability” of performance and personality traits may impact “other” ratings. Another example would be an examination of ratings from instructors who provided training to the employees, which may further inform how personality interacts with rating congruence in a training context. Rating tendencies

in training research have been overlooked and may provide valuable information on how to tailor training and feedback interventions.

Future studies should also carefully assess the best statistical analytical methods for their research, such as polynomial regression analysis with agreement categories, with and between analysis (WABA), and other various statistical approaches used to create an index of self-other agreement (Atwater & Yammarino, 1997; Edwards, 1995; Fleenor et al., 2010). One example that may provide further nuance to how personality influences rating congruence would be examining possible interactions between personality facets. Another example would be to take a person-centric or profile approach to personality traits and their links to rating tendencies. The current study acknowledges the limited scope of interpretability with the chosen analytical method of multivariate regression. However, while WABA and polynomial regression may be desirable techniques for examining self-other agreement among multiple sources, these analyses were not appropriate to answer the current research question. Future research should continue to examine this stream of research using a variety of appropriate statistical analyses.

In conclusion, the current study's findings may serve as a beginning for understanding narrow personality facets as antecedents of rating congruence. Research and practice in the areas of leadership training and development associated with talent management applications of 360-degree feedback systems

will benefit from the continued investigation of the relationship between personality and rating agreement.

## References

- Abramson, L. Y., & Andrews, D. E. (1982). Cognitive models of depression: Implications for sex differences in vulnerability to depression. *International Journal of Mental Health, 11*(1–2), 77–94.  
<https://doi.org/10.1080/00207411.1982.11448907>
- Alloy, L. B., & Abramson, L. Y. (1982). Learned helplessness, depression, and the illusion of control. *Journal of Personality and Social Psychology, 42*(6), 1114–1126. <https://doi.org/10.1037/0022-3514.42.6.1114>
- Alloy, L. B., Abramson, L. Y., & Viscusi, D. (1981). Induced mood and the illusion of control. *Journal of Personality and Social Psychology, 41*(6), 1129–1140. <https://doi.org/10.1037/0022-3514.41.6.1129>
- Antonioni, D. (1996). Designing an effective 360-degree appraisal feedback process. *Organizational Dynamics, 25*(2), 24–39.  
[https://doi.org/10.1016/S0090-2616\(96\)90023-6](https://doi.org/10.1016/S0090-2616(96)90023-6)
- Antonioni, D., & Park, H. (2001). The effects of personality similarity on peer ratings of contextual work behaviors. *Personnel Psychology, 54*(2), 331–360. <https://doi.org/10.1111/j.1744-6570.2001.tb00095.x>
- Ashford, S. J. (1989). Self-assessments in organizations-a literature-review and integrative model. In *Research in organizational behavior* (Vol. 11, pp. 133–174). JAI Press.

- Atwater, L. E., Brett, J. F., & Charles, A. C. (2007). Multisource feedback: Lessons learned and implications for practice. *Human Resource Management, 46*(2), 285–307. <https://doi.org/10/fqz9qv>
- Atwater, L. E., Ostroff, C., Yammarino, F. J., & Fleenor, J. W. (1998). Self–other agreement: Does it really matter? *Personnel Psychology, 51*(3), 577–598. <https://doi.org/10.1111/j.1744-6570.1998.tb00252.x>
- Atwater, L. E., & Yammarino, F. J. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology, 45*(1), 141. <https://doi.org/10.1111/j.1744-6570.1992.tb00848.x>
- Atwater, L. E., & Yammarino, F. J. (1997). Self-other rating agreement: A review and model. In *Research in personnel and human resources management, Vol. 15* (pp. 121–174). Elsevier Science/JAI Press.
- Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917–1992. *Journal of Applied Psychology, 77*(6), 836–874. <https://doi.org/10.1037/0021-9010.77.6.836>
- Bailey, C., & Fletcher, C. (2002). The impact of multiple source feedback on management development: Findings from a longitudinal study. *Journal of Organizational Behavior, 23*(7), 853–867. <https://doi.org/10.1002/job.167>
- Bailey, K. G., & Lazar, J. (1976). Accuracy of self-ratings of intelligence as a function of sex and level of ability in college students. *Journal of Genetic*

*Psychology*, 129(2), 279–290.

<https://doi.org/10.1080/00221325.1976.10534039>

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next? *International Journal of Selection & Assessment*, 9(1/2), 9. <https://doi.org/10.1111/1468-2389.00160>

Bass, B. M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology*, 40(4), 437–454. <https://doi.org/10.1111/j.1464-0597.1991.tb01002.x>

Bayroff, A. G., Haggerty, H. R., & Rundquist, E. A. (1954). Validity of ratings as related to rating techniques and conditions. *Personnel Psychology*, 7(1), 93–113. <https://doi.org/10.1111/j.1744-6570.1954.tb02262.x>

Beehr, T. A., Ivanitskaya, L., Hansen, C. P., Erofeev, D., & Gudanowski, D. M. (2001). Evaluation of 360 degree feedback ratings: Relationships with each other and with performance and selection predictors. *Journal of Organizational Behavior*, 22(7), 775–788. <https://doi.org/10.1002/job.113>

Bell, S. T., & Arthur Jr., W. (2008). Feedback acceptance in developmental assessment centers: The role of feedback message, participant personality, and affective response to the feedback session. *Journal of Organizational Behavior*, 29(5), 681–703. <https://doi.org/10.1002/job.525>

- Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74(3), 183–200.  
<https://doi.org/10.1037/h0024835>
- Bergner, S., Davda, A., Culpin, V., & Rybnicek, R. (2016). Who overrates, who underrates? Personality and its link to self–other agreement of leadership effectiveness. *Journal of Leadership & Organizational Studies*, 23(3), 335–354. <https://doi.org/10.1177/1548051815621256>
- Boyle, G., Matthews, G., & Saklofske, D. (2008). The California Psychological Inventory – 434- and 260-item Editions. In *The SAGE Handbook of Personality Theory and Assessment: Volume 2—Personality Measurement and Testing*. SAGE Publications Ltd.  
<https://doi.org/10.4135/9781849200479>
- Bracken, D. W., Rose, D. S., & Church, A. H. (2016). The evolution and devolution of 360° feedback. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 9(4), 761–794.  
<https://doi.org/10.1017/iop.2016.93>
- Bracken, D. W., Timmreck, C. W., & Church, A. H. (2001). *The Handbook of Multisource Feedback*. John Wiley & Sons.
- Brett, J. F., & Atwater, L. E. (2001). 360° feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology*, 86(5), 930–942.  
<https://doi.org/10/cnz8h8>

- Brief, A. P., Aldag, R. J., & Van Sell, M. (1977). Moderators of the relationships between self and superior evaluations of job performance. *Journal of Occupational Psychology*, 50(2), 129–134. <https://doi.org/10.1111/j.2044-8325.1977.tb00366.x>
- Brook, J. A. (1991). The link between self-esteem and work/nonwork perceptions and attitudes. *Applied Psychology: An International Review*, 40(3), 269–280. <https://doi.org/10.1111/j.1464-0597.1991.tb00990.x>
- Brutus, S., Fleenor, J. W., & McCauley, C. D. (1999). Demographic and personality predictors of congruence in multi-source ratings. *Journal of Management Development*, 18(5), 417–435. <https://doi.org/10.1108/02621719910273569>
- Burusic, J., & Ribar, M. (2014). The moderating role of self-presentation tactics: Judgments of personality traits and self-presentation of others based on a limited amount of information. *Swiss Journal of Psychology*, 73(4), 235–242. <https://doi.org/10.1024/1421-0185/a000142>
- Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice*, 31, 141–154. <https://doi.org/10.1037/0735-7028.31.2.141>



- Campbell, D. J., & Lee, C. (1988). Self-appraisal in performance evaluation: Development versus evaluation. *Academy of Management Review*, 13(2), 302–314. <https://doi.org/10.5465/amr.1988.4306896>
- Carter, M. J., & Fuller, C. (2015). Symbolic interactionism. *Sociopedia*. <https://doi.org/10.1177/205684601561>
- Cawley, B. D., Keeping, L. M., & Levy, P. E. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of field investigations. *Journal of Applied Psychology*, 83(4), 615–633. <http://dx.doi.org.portal.lib.fit.edu/10.1037/0021-9010.83.4.615>
- Church, A. H., Bracken, D. W., Fleenor, J. W., & Rose, D. S. (2019). *Handbook of Strategic 360 Feedback*. Oxford University Press.
- Church, A. H., & Rotolo, C. T. (2013). How are top companies assessing their high-potentials and senior executives? A talent management benchmark study. *Consulting Psychology Journal: Practice and Research*, 65, 199–223. <https://doi.org/10.1037/a0034381>
- Cobb, C. L., Xie, D., & Maragakis, A. (2020). California Psychological Inventory (CPI). In *The Wiley Encyclopedia of Personality and Individual Differences* (pp. 273–277). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119547167.ch117>
- Colarelli, Meyer & Associates, Inc. (2012). *Insight 360 Validation Report*. <https://cmaconsult.com/>

- Conway, J. M., & Huffcutt, A. I. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance, 10*(4), 331–360.  
[https://doi.org/10.1207/s15327043hup1004\\_2](https://doi.org/10.1207/s15327043hup1004_2)
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98–104.  
<http://dx.doi.org.portal.lib.fit.edu/10.1037/0021-9010.78.1.98>
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*(4), 349–354. <https://doi.org/10.1037/h0047358>
- Day, D. V., Shleicher, D. J., Unckless, A. L., & Hiller, N. J. (2002). Self-monitoring personality at work: A meta-analytic investigation of construct validity. *Journal of Applied Psychology, 87*(2), 390–401.  
<https://doi.org/10.1037/0021-9010.87.2.390>
- Dunnette, M. D. (1993). My hammer or your hammer? *Human Resource Management, 32*(2/3), 373–384. <https://doi.org/10.1002/hrm.3930320212>
- Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes, 64*(3), 307–324.  
<https://doi.org/10.1006/obhd.1995.1108>

- Evans-Reber, K. (2020, February 19). Council Post: Here's What Can Happen When Companies Get Rid Of Performance Reviews. *Forbes*.  
<https://www.forbes.com/sites/forbeshumanresourcescouncil/2020/02/19/heres-what-can-happen-when-companies-get-rid-of-performance-reviews/>
- Facteau, C. L., Facteau, J. D., Schoel, L. C., Russell, J. E. A., & Poteet, M. L. (1998). Reactions of leaders to 360-degree feedback from subordinates and peers. *The Leadership Quarterly*, *9*(4), 427–448.  
[https://doi.org/10.1016/S1048-9843\(98\)90010-8](https://doi.org/10.1016/S1048-9843(98)90010-8)
- Farh, J., & Werbel, J. D. (1986). Effects of purpose of the appraisal and expectation of validation on self-appraisal leniency. *Journal of Applied Psychology*, *71*(3), 527–529. <https://doi.org/10.1037/0021-9010.71.3.527>
- Farh, J.-L., & Dobbins, G. H. (1989). Effects of self-esteem on leniency bias in self-reports of. *Personnel Psychology*, *42*(4), 835.  
<https://doi.org/10.1111/j.1744-6570.1989.tb00677.x>
- Farr, J. L., & Newman, D. A. (2001). Rater selection: Sources of feedback. In *The Handbook of Multisource Feedback* (pp. 96–113). John Wiley & Sons.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, *116*(3), 429–456. <https://doi.org/10.1037/0033-2909.116.3.429>

- Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology, 43*(4), 522–527. <https://doi.org/10.1037/h0076760>
- Ferris, G. R., & Judge, T. A. (1991). Personnel/human resources management: A political influence perspective. *Journal of Management, 17*(2), 447. <https://doi.org/10.1177/014920639101700208>
- Fleenor, J. W., McCauley, C. D., & Brutus, S. (1996). Self-other rating agreement and leader effectiveness. *The Leadership Quarterly, 7*(4), 487–506. [https://doi.org/10.1016/S1048-9843\(96\)90003-X](https://doi.org/10.1016/S1048-9843(96)90003-X)
- Fleenor, J. W., Smither, J. W., Atwater, L. E., Braddy, P. W., & Sturm, R. E. (2010). Self–other rating agreement in leadership: A review. *The Leadership Quarterly, 21*(6), 1005–1034. <https://doi.org/10.1016/j.leaqua.2010.10.006>
- Fletcher, C. (2014). Using 360-Degree Feedback as a Development Tool. In *The Wiley Blackwell Handbook of the Psychology of Training, Development, and Performance Improvement* (pp. 486–502). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118736982.ch24>
- Funder, D. C. (2006). Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality, 40*(1), 21–34. <https://doi.org/10.1016/j.jrp.2005.08.003>

- Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52*, 409–418. <https://doi.org/10.1037/0022-3514.52.2.409>
- Furnham, A., Moutafi, J., & Chamorro-Premuzic, T. (2005). Personality and intelligence: Gender, the Big Five, self-estimated and psychometric intelligence. *International Journal of Selection and Assessment, 13*(1), 11–24. <https://doi.org/10.1111/j.0965-075X.2005.00296.x>
- Gentry, W. A., Ekelund, B. Z., Hannum, K. M., & de Jong, A. (2007). A study of the discrepancy between self- and observer-ratings on managerial derailment characteristics of European managers. *European Journal of Work and Organizational Psychology, 16*(3), 295–325. <https://doi.org/10.1080/13594320701394188>
- Goffin, R. D., & Anderson, D. W. (2007). The self-rater's personality and self-other disagreement in multi-source performance ratings: Is disagreement healthy? *Journal of Managerial Psychology, 22*(3), 271–289. <https://doi.org/10.1108/02683940710733098>
- Gough, H. G. (2009). *CPI Additional Interpretation Guide*. CPP, Inc.
- Gough, H. G., & Bradley, P. (2005). *CPI 260 Manual*. CPP, Inc.
- Hall, J. L., Leiaecker, J. K., & DiMarco, C. (1996). What we know about upward appraisals of management: Facilitating the future use of UPAs. *Human*

*Resource Development Quarterly*, 7(3), 209–226.

<https://doi.org/10.1002/hrdq.3920070303>

Halverson, S. K., Tonidandel, S., Barlow, C., & Dipboye, R. L. (2002). *Self-other agreement on a 360-degree leadership evaluation*. 23.

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

<https://doi.org/10.1111/j.1744-6570.1988.tb00631.x>

Hjelle, L. A. (1969). Personality characteristics associated with interpersonal perception accuracy. *Journal of Counseling Psychology*, 16(6), 579–581.

<https://doi.org/10.1037/h0028439>

Hoffman, G. J. (1923). An experiment in self-estimation. *The Journal of Abnormal Psychology and Social Psychology*, 18(1), 43–49.

<https://doi.org/10.1037/h0064542>

Ingram, R. E. (1989). Unique and Shared Cognitive Factors in Social Anxiety and Depression: Automatic Thinking and Self-Appraisal. *Journal of Social and Clinical Psychology*, 8(2), 198–208.

<https://doi.org/10.1521/jscp.1989.8.2.198>

Jackson, D. J. R., Stillman, J. A., Burke, S., & Englert, P. (2007). Self versus assessor ratings and their classification in assessment centres: Profiling the self-rater. *New Zealand Journal of Psychology*, 36(2), 93–99.

- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, *61*(4), 521–551. <https://doi.org/10.1111/j.1467-6494.1993.tb00781.x>
- Judge, T. A., LePine, J. A., & Rich, B. L. (2006). Loving yourself abundantly: Relationship of the narcissistic personality to self- and other perceptions of workplace deviance, leadership, and task and contextual performance. *Journal of Applied Psychology*, *91*, 762–776. <https://doi.org/10.1037/0021-9010.91.4.762>
- Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, *31*(1), 457–501. <https://doi.org/10.1146/annurev.ps.31.020180.002325>
- Kenny, D. A. (1994). Interpersonal perception: A social relations analysis. In *Interpersonal perception: A social relations analysis* (p. Chapter xvii, 270 Pages). Guilford Press (New York, NY, US). <https://www.proquest.com/psycinfo/docview/618584271/D13C9885F00F4347PQ/3>
- Kernis, M. H., Grannemann, B. D., & Mathis, L. C. (1991). Stability of self-esteem as a moderator of the relation between level of self-esteem and depression. *Journal of Personality and Social Psychology*, *61*(1), 80–84. <https://doi.org/10.1037/0022-3514.61.1.80>

- Kim, K. Y., Atwater, L., Patel, P. C., & Smither, J. W. (2016). Multisource feedback, human capital, and the financial performance of organizations. *Journal of Applied Psychology, 101*(11), 1569–1584.  
<https://doi.org/10/f9b7fr>
- Kinicki, A. J., Prussia, G. E., Bin (Joshua) Wu, & McKee-Ryan, F. M. (2004). A covariance structure analysis of employees' response to performance feedback. *Journal of Applied Psychology, 89*(6), 1057–1069.  
<https://doi.org/10.1037/0021-9010.89.6.1057>
- Klimoski, R. J., & London, M. (1974). Role of the rater in performance appraisal. *Journal of Applied Psychology, 59*(4), 445–451.  
<https://doi.org/10.1037/h0037332>
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.  
<https://doi.org/10.1037/0033-2909.119.2.254>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134.  
<https://doi.org/10.1037/0022-3514.77.6.1121>



- Levine, E. L., Flory, A., & Ash, R. A. (1977). Self-assessment in personnel selection. *Journal of Applied Psychology, 62*(4), 428–435.  
<https://doi.org/10.1037/0021-9010.62.4.428>
- Levy, P. E. (1993). Self-appraisal and attributions: A test of a model. *Journal of Management, 19*(1), 51–63. <https://doi.org/10.1177/014920639301900104>
- Lewis, M. A., & Neighbors, C. (2005). Self-determination and the use of self-presentation strategies. *The Journal of Social Psychology, 145*(4), 469–489.  
<https://doi.org/10.3200/SOCP.145.4.469-490>
- London, M. (1994). Interpersonal insight in organizations: Cognitive models for human resource development. *Human Resource Management Review, 4*(4), 311–332. [https://doi.org/10.1016/1053-4822\(94\)90017-5](https://doi.org/10.1016/1053-4822(94)90017-5)
- London, M., & Smither, J. W. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review, 12*(1), 81–100. <https://doi.org/10/c9kbvn>
- Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*, 280–296.  
<https://doi.org/10.1037/0021-9010.67.3.280>
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology, 43*(2), 293–303. <https://doi.org/10.1037/0022-3514.43.2.293>

- McCrae, R. R., Costa Jr., P. T., & Piedmont, R. L. (1993). Folk Concepts, Natural Language, and Psychological Constructs: The California Psychological Inventory and the Five-Factor Model. *Journal of Personality*, *61*(1), 1–26. <https://doi.org/10.1111/j.1467-6494.1993.tb00276.x>
- McKee, R. A., Lee, Y., Atwater, L., & Antonakis, J. (2018). Effects of personality and gender on self–other agreement in ratings of leadership. *Journal of Occupational and Organizational Psychology*, *91*(2), 285–315. <https://doi.org/10.1111/joop.12209>
- Miller, J. S., & Cardy, R. L. (2000). Self-monitoring and performance appraisal: Rating outcomes in project teams. *Journal of Organizational Behavior*, *21*(6), 609. [https://doi.org/10.1002/1099-1379\(200009\)21:6<609::AID-JOB42>3.0.CO;2-K](https://doi.org/10.1002/1099-1379(200009)21:6<609::AID-JOB42>3.0.CO;2-K)
- Moshavl, D., Brown, F. W., & Dodd, N. G. (2003). Leader self-awareness and its relationship to subordinate attitudes and performance. *Leadership & Organization Development Journal*, *24*(7/8), 407–418. <https://doi.org/10.1108/01437730310498622>
- Needham, S. M., & Jacobson, J. A. (2020). Schemas, Theory of. In *The Wiley Encyclopedia of Personality and Individual Differences* (pp. 363–367). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118970843.ch60>
- Oregon Research Institute. (2022, February 21). *CPI Comparison Table*. International Personality Item Pool. <https://ipip.ori.org/newCPITable.htm>

- Ostroff, C., Atwater, L. E., & Feinberg, B. J. (2004). Understanding self-other agreement: A look at rater and ratee characteristics, context, and outcomes. *Personnel Psychology, 57*(2), 333–375. <https://doi.org/10.1111/j.1744-6570.2004.tb02494.x>
- Park, B., & Judd, C. M. (1989). Agreement on initial impressions: Differences due to perceivers, trait dimensions, and target behaviors. *Journal of Personality and Social Psychology, 56*(4), 493–505. <https://doi.org/10.1037/0022-3514.56.4.493>
- Pyant, C. T., & Yanico, B. J. (1991). Relationship of racial identity and gender-role attitudes to Black women's psychological well-being. *Journal of Counseling Psychology, 38*(3), 315–322. <https://doi.org/10.1037/0022-0167.38.3.315>
- Robins, R. W., Tracy, J. L., Trzesniewski, K., Potter, J., & Gosling, S. D. (2001). Personality correlates of self-esteem. *Journal of Research in Personality, 35*(4), 463–482. <https://doi.org/10.1006/jrpe.2001.2324>
- Roethlisberger, F. J., & Dickson, W. J. (2003). *Management and the Worker*. Psychology Press.
- Rothstein, H. R. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology, 75*(3), 322–327. <https://doi.org/10.1037/0021-9010.75.3.322>

- Ruble, D. N., & Flett, G. L. (1988). Conflicting Goals in Self-Evaluative Information Seeking: Developmental and Ability Level Analyses. *Child Development, 59*(1), 97–106. <https://doi.org/10.2307/1130392>
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*(1), 141. <https://doi.org/10.1146/annurev.psych.52.1.141>
- Ryff, C. D., & Singer, B. (1998). The Contours of Positive Human Health. *Psychological Inquiry, 9*(1), 1. [https://doi.org/10.1207/s15327965pli0901\\_1](https://doi.org/10.1207/s15327965pli0901_1)
- Sala, F. (2003). Executive blind spots discrepancies between self- and other-ratings. *Consulting Psychology Journal: Practice & Research, 55*(4), 222–229. <https://doi.org/10.1037/1061-4087.55.4.222>
- Schlenker, B. R., Weigold, M. F., & Hallam, J. R. (1990). Self-serving attributions in social context: Effects of self-esteem and social pressure. *Journal of Personality and Social Psychology, 58*(5), 855–863. <https://doi.org/10.1037/0022-3514.58.5.855>
- Sinha, N. (2004). Self-other rating discrepancy and personality antecedents. In *Dissertation Abstracts International: Section B: The Sciences and Engineering* (Vol. 65, Issues 3-B, p. 1587). Florida International University.
- Sinha, N., Mesmer-Magnus, J., & Viswesvaran, C. (2012). Personality antecedents of self-other rating discrepancy. *Revista de Psicología Del Trabajo y de Las Organizaciones, 28*(1), 3–14. <https://doi.org/10.5093/tr2012a1>

- Smith, T. A., Boulamatsi, A., Dimotakis, N., Tepper, B. J., Runnalls, B. A., Reina, C. S., & Lucianetti, L. (2022). “How dare you?!”: A self-verification perspective on how performance influences the effects of abusive supervision on job embeddedness and subsequent turnover. *Personnel Psychology, 75*(3), 645–674. <https://doi.org/10.1111/peps.12494>
- Smither, J. W., London, M., & Reilly, R. R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology, 58*(1), 33–66. [https://doi.org/10.1111/j.1744-6570.2005.514\\_1.x](https://doi.org/10.1111/j.1744-6570.2005.514_1.x)
- Smither, J. W., London, M., Vasilopoulos, N. L., Reilly, R. R., Millsap, R. E., & Salvemini, N. (1995). An examination of the effects of an upward feedback program over time. *Personnel Psychology, 48*(1), 1–34. <https://doi.org/10/fdsq3m>
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*(4), 526–537. <https://doi.org/10.1037/h0037039>
- Sorrentino, R. M., & Hewitt, E. C. (1984). The uncertainty-reducing properties of achievement tasks revisited. *Journal of Personality and Social Psychology, 47*(4), 884–899. <https://doi.org/10.1037/0022-3514.47.4.884>

- Swann, W. B., Pelham, B. W., & Krull, D. S. (1989). Agreeable fancy or disagreeable truth? Reconciling self-enhancement and self-verification. *Journal of Personality and Social Psychology*, *57*(5), 782–791.  
<https://doi.org/10.1037/0022-3514.57.5.782>
- Szell, S., & Henderson, R. (1997). The impact of self-supervisor/subordinate performance rating agreement on subordinates' job satisfaction and organisational commitment. *Journal of Applied Social Behaviour*, *3*, 25–37.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*(2), 193–210. <https://doi.org/10.1037/0033-2909.103.2.193>
- Tetlock, P. E. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quarterly*, *46*(4), 285–292.  
<https://doi.org/10.2307/3033716>
- Tetlock, P. E., & Boettger, R. (1989). Accountability: A social magnifier of the dilution effect. *Journal of Personality and Social Psychology*, *57*(3), 388–398. <https://doi.org/10.1037/0022-3514.57.3.388>
- Thornton, G. C. (1980). Psychometric properties of self-appraisals of job performance. *Personnel Psychology*, *33*(2), 263–271.  
<https://doi.org/10.1111/j.1744-6570.1980.tb02348.x>

- Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*(3), 239–257.  
<https://doi.org/10.1037/0033-295X.93.3.239>
- Tunnell, G. (1980). Intraindividual consistency in personality assessment: The effect of self-monitoring. *Journal of Personality*, *48*(2), 220–232.  
<https://doi.org/10.1111/j.1467-6494.1980.tb00829.x>
- Van Velsor, E., Taylor, S., & Leslie, J. B. (1993). An examination of the relationships among self-perception accuracy, self-awareness, gender, and leader effectiveness. *Human Resource Management*, *32*(2/3), 249–263.  
<https://doi.org/10.1002/hrm.3930320205>
- Vaughan-Johnston, T. I., MacGregor, K. E., Fabrigar, L. R., Evraire, L. E., & Wasylkiw, L. (2021). Extraversion as a moderator of the efficacy of self-esteem maintenance strategies. *Personality and Social Psychology Bulletin*, *47*(1), 131–145. <https://doi.org/10.1177/0146167220921713>
- Vecchio, R. P., & Anderson, R. J. (2009). Agreement in self–other ratings of leader effectiveness: The role of demographics and personality. *International Journal of Selection and Assessment*, *17*(2), 165–179.  
<https://doi.org/10.1111/j.1468-2389.2009.00460.x>
- Visser, B. A., Ashton, M. C., & Vernon, P. A. (2008). What makes you think you're so smart? Measured abilities, personality, and sex differences in

- relation to self-estimates of multiple intelligences. *Journal of Individual Differences*, 29, 35–44. <https://doi.org/10.1027/1614-0001.29.1.35>
- Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, 59(2), 197–210. <https://doi.org/10.1177/00131649921969802>
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81(5), 557–574. <https://doi.org/10.1037/0021-9010.81.5.557>
- Ward, P. (1997). *360-degree Feedback*. CIPD Publishing.
- Weinstein, E. A. (1969). The development of interpersonal competence. In *Handbook of Socialization Theory and Research* (pp. 753–775). Guilford Press.
- Wells, L. E., & Sweeney, P. D. (1986). A Test of Three Models of Bias in Self-Assessment. *Social Psychology Quarterly*, 49(1), 1–10. <https://doi.org/10.2307/2786852>
- Wohlers, A. J., & London, M. (1989). Ratings of managerial characteristics: Evaluation difficulty, co-worker agreement, and self-awareness. *Personnel Psychology*, 42(2), 235–261. <https://doi.org/10.1111/j.1744-6570.1989.tb00656.x>
- Wood, J. V., Giordano-Beech, M., Taylor, K. L., Michela, J. L., & Gaus, V. (1994). Strategies of social comparison among people with low self-esteem: Self-



protection and self-enhancement. *Journal of Personality and Social Psychology*, 67(4), 713–731. <https://doi.org/10.1037/0022-3514.67.4.713>

Wymer, W. E., & Penner, L. A. (1985). Moderator variables and different types of predictability: Do you have a match? *Journal of Personality and Social Psychology*, 49(4), 1002–1015. <https://doi.org/10.1037/0022-3514.49.4.1002>

Yammarino, F. J., & Atwater, L. E. (1993). Understanding self-perception accuracy: Implications for human resource management. *Human Resource Management*, 32(2–3), 231–247. <https://doi.org/10.1002/hrm.3930320204>

Yammarino, F. J., & Atwater, L. E. (1997). Do managers see themselves as others see them? Implications of self-other rating agreement for human resources management. *Organizational Dynamics*, 25(4), 35–44. [https://doi.org/10.1016/S0090-2616\(97\)90035-8](https://doi.org/10.1016/S0090-2616(97)90035-8)

## Tables

**Table 1: CPI 260 Scales and Their Purposes**

<b>Scales</b>	<b>Intentions</b>
<b>Dealing with others</b>	
Dominance (Do)	To assess prosocial interpersonal dominance, strength of will, and perseverance in pursuit of goals
Capacity for Status (Cs)	To measure personal qualities that are associated with and that lead to high social status, including ambition and self-confidence
Sociability (Sy)	To identify people who are outgoing and socially affiliative, and who enjoy social participation
Social Presence (Sp)	To identify people who are self-assured, comfortable being the center of attention, and socially adroit
Self-acceptance (Sa)	To identify people with high self-esteem, a strong sense of personal worth, and optimism
Independence (In)	To assess the twin elements of psychological strength and interpersonal detachment, including self-sufficiency and self-direction
Empathy (Em)	To identify people with a talent for understanding how others feel and think, and who display warmth and tactfulness in their dealings with others
<b>Self-management</b>	
Responsibility (Re)	To identify people who are aware of societal rules, and who can and do comply with them when this is appropriate

Social Conformity (So)	To assess the degree to which societal norms have been internalized and become autonomously operational within the individual
Self-control (Sc)	To assess a continuum going from under control and expressiveness at one pole to overcontrol and suppression of affect at the other
Good Impression (GI)	First, for very high scores, to identify overly strong attempts to create a favorable impression; and second, to identify people whose style of self-presentation emphasizes ingratiation and compliance
Communality (Cm)	To assess a continuum going from erratic or random answering at one pole to close agreement with ordinary beliefs and conventions at the other
Well-being (Wb)	To assess feelings of physical and psychological well-being
Tolerance (To)	To assess attitudes of tolerance, forbearance, and respect for others, stemming from ethical convictions about the worth of all people
<b>Motivations and Thinking Style</b>	
Achievement via Conformance (Ac)	To assess achievement potential in well-defined and structured situations, joined to a general desire to do well
Achievement via Independence (Ai)	To assess achievement potential in open, minimally defined situations in which ingenuity and initiative are required for successful performance
Conceptual Fluency (Cf)	To identify people who deal easily with abstract and complex concepts, and who believe in their own talent
<b>Personal Characteristics</b>	
Insightfulness (Is)	To identify people who can think analytically about themselves and others, who can see beyond surface cues, and who are aware of subtle meanings

Flexibility (Fx)	To assess a continuum going from resistance to change and dislike of uncertainty at one pole to a liking for change and innovation at the other
Sensitivity (Sn)	To assess a continuum going from tough-minded practicality and relative uninterest in personal feelings at one pole to sensitivity, solicitude for others, and a sense of own vulnerability at the other
<b>Work-Related Measures</b>	
Managerial Potential (Mp)	To identify people with an interest in management and who have effective interpersonal skills and good judgment
Work Orientation (Wo)	To identify people with a dutiful work ethic, a strong sense of commitment to their job, and little need for overt recognition
Creative Temperament (Ct)	To identify people of an imaginative, creative temperament, with both the need and potential for visualizing new and different ways of doing things
Leadership (Lp)	To identify people who have good leadership skills, who aspire to positions of leadership, and who will be accepted as leaders by others
Amicability (Ami)	To identify people who are amicable, friendly, and considerate of others, who try to avoid conflicts, and who seldom become angry or irritated
Law Enforcement Orientation (Leo)	To identify people who view law enforcement and societal rules favorably, who believe punishment for violation of such rules is deserved, and who are well-suited for work in the law enforcement field

**Note:** From “CPI 260 Manual” by Gough, H.G. and Bradley P., 2005, pg. 7. Copyright 2005 by CPP, Inc.

**Table 2: 360-degree Instrument Scale Descriptions**

<b><i>Dimension</i></b>	<b><i>Definition</i></b>
<b>Analysis &amp; Decision Making</b>	Making effective decisions and solving problems in complex or ambiguous situations by gathering, diagnosing, and judiciously analyzing the information about the situation and environment in order to identify and evaluate options and to select the best course of action.
<b>Responsibility &amp; Achievement</b>	Demonstrating initiative, commitment to excellence, and effective self-management skills, including responsibility, dependability, planning and organization, detail-orientation, and follow through.
<b>Relationships &amp; Teaming</b>	Effectively building relationships with individuals and teams across the organization by being inclusive, considerate, and responsive to the needs of others; by communicating effectively, collaborating with others, and sharing resources; and by being receptive to feedback.
<b>Managing &amp; Developing</b>	Managing the work of others by providing direction, structure, and clear expectations; maintaining an open flow of communication and a sense of urgency to drive results and hold others accountable; and developing others by providing timely and relevant feedback and opportunities for development.
<b>Leading &amp; Influencing</b>	Inspiring and influencing by communicating a compelling vision of the future, conveying an executive presence (e.g., confidence, poise, connecting with others), and being sufficiently agile and self-assured to effectively lead others.

*Note:* From “Insight 360 Validation Report” by Colarelli, Meyer & Associates, Inc., 2012, pg. 3.

**Table 3: Personality, Performance, and Rating Tendency Correlations**

*Personality, Performance, and Rating Tendency Correlations*

	M	SD	1	2	3	4	5	6	7	8	9	10
1. Dominance	62.55	6.38	--									
2. Sociability	56.64	7.28	.73**	--								
3. Well-Being	58.66	4.95	.31**	.20**	--							
4. Empathy	63.15	7.51	.47**	.52**	.28**	--						
5. Good Impression	61.85	8.79	.23**	.19**	.55**	.22**	--					
6. Overall Perf - Self	7.33	.636	.33**	.29**	.31**	.15*	.35**	--				
7. Overall Perf - Supe	7.33	.686	.10	.04	.10	.00	.09	.12*	--			
8. Overall Perf - Peer	7.44	.567	-.01	.07	.12	.04	.17**	.16*	.29**	--		
9. Overall Perf - DR	7.69	.545	-.01	.02	.01	.00	.04	.18**	.18*	.31**	--	
10. Overall Perf - Other	7.46	.49	.03	.05	.11	.00	.14*	.17**	.79**	.74**	.69**	--
11. Overrater Supe	.279	.449	.09	.10	.09	.06	.16**	.45**	-.57**	-.11	-.03	-.42**
12. Underrater Supe	.333	.472	-.14*	-.16**	-.12*	-.04	-.12*	-.51**	.48**	.03	.00	.29**
13. Agree Supe	.333	.472	.06	.05	.01	-.03	-.08	.03	.06	.02	.00	.04
14. Overrater Peer	.194	.396	.13*	.05	.05	-.02	.14*	.43**	-.03	-.46**	.02	-.22**
15. Underrater Peer	.330	.471	-.24**	-.17**	-.06	-.09	-.12*	-.53**	.00	.44**	.01	.15**
16. Agree Peer	.296	.457	.06	.02	.00	.04	-.02	.05	.00	-.04	-.02	-.01
17. Overrater Sub	.109	.312	.21**	.13*	.12*	.10	.23**	.37**	-.03	-.04	-.46**	-.18**
18. Underrater Sub	.391	.488	-.22**	-.22**	-.09	-.04	-.15**	-.46**	.09	.05	.39**	.27**
19. Agree Sub	.242	.429	.15**	.19**	.12*	.04	.06	.28**	.03	.06	-.06	.03
20. Overrater Other	.248	.432	.20**	.16**	.12*	.07	.17**	.57**	-.32**	-.29**	-.20**	-.39**
21. Underrater Other	.391	.489	-.22**	-.22**	-.18**	-.07	-.15*	-.61**	.29**	.22**	.17*	.36**
22. Agree Other	.343	.476	.09	.12*	.11	.00	.00	.29**	-.07	-.04	.17*	.04

Note: \*  $p < .05$ ; \*\*  $p < .01$

**Table 4: Multivariate Regression Results**

Multivariate Regression Results

Personality Trait	Self-Supe		Self-Peer		Self-Direct Report (DR)		Self-Other	
	Self (eta)	Supe (eta)	Self (eta)	Peer (eta)	Self (eta)	DR (eta)	Self (eta)	Other (eta)
Dominance	<b>.033*** (.107)</b>	<b>.011</b>	<b>.032*** (.103)</b>	<b>-.002</b>	<b>.035*** (.123)</b>	<b>-.001</b>	<b>.033*** (.107)</b>	<b>.003</b>
Sociability	<b>.025*** (.085)</b>	<b>.004</b>	<b>.024*** (.079)</b>	<b>.005</b>	<b>.030*** (.114)</b>	<b>.002</b>	<b>.025*** (.085)</b>	<b>.004</b>
Well Being	<b>.039*** (.096)</b>	<b>.014</b>	<b>.038*** (.091)</b>	<b>.014</b>	<b>.035*** (.064)</b>	<b>.002</b>	<b>.040*** (.096)</b>	<b>.011</b>
Empathy	.012	.000	.010	.004	.011	.001	<b>.012** (.021)</b>	<b>.000</b>
Good Impression	<b>.024*** (.116)</b>	<b>.007</b>	<b>.028*** (.145)</b>	<b>.011* (.028)</b>	<b>.025*** (.109)</b>	<b>.003</b>	<b>.025*** (.123)</b>	<b>.008* (.020)</b>

Note: \* p < .05, \*\* p < .01, \*\*\* p < .001; Bolded results indicate significant Wilks'  $\Lambda$ ; (eta) indicates the proportion of variance explained and is only reported with significant predictors

**Table 5: Exploratory Correlations**

*Personality Facets and Rating Tendencies Exploratory Correlations*

	Do	Cs	Sy	Sp	Sa	In	Em	Re	So	Sc	Gi	Cm	Wb	To	Ac	Ai	Cf	Is	Fx	Sn	Mp	Wo	Ct	Lp	Ami	Leo
Ovrrater Supe	.09	.11	.10	.00	.07	.05	.06	.09	.12*	.09	.16**	.06	.09	.10	.07	.04	.04	.09	-.02	-.08	.10	.12*	-.01	.10	.08	.11
Underrater Supe	-.14*	-.13*	-.16**	-.08	-.13*	-.12*	-.04	-.07	-.09	-.04	-.12*	-.06	-.12*	-.15*	-.08	-.10	-.14*	-.21**	.02	.13*	-.19**	-.07	-.06	-.18**	-.08	-.04
Agree Supe	.06	.03	.05	.10	.05	.06	0.03	-.04	.01	-.07	-.08	-.01	.01	.01	-.02	.00	.06	.14*	.01	-.04	.05	-.07	.06	.06	-.02	-.05
Ovrrater Peer	.13*	.03	.05	-.11	.05	.13*	0.02	.12*	.14*	.12*	.14*	.01	.05	.12*	.05	-.01	.04	.05	-.09	-.15**	.13*	.05	-.01	.13*	.07	.14*
Underrater Peer	-.24**	-.19**	-.17**	-.08	-.17**	-.22**	-.09	-.10	-.14*	-.03	-.12*	.01	-.06	-.11	-.07	-.02	-.14*	-.04	.09	.14*	-.22**	-.07	-.07	-.25**	-.09	-.03
Agree Peer	.06	.06	.02	.06	.07	.14*	.04	.02	-.01	-.04	-.02	.01	.00	-.03	.00	.01	.04	.02	.01	.04	.03	-.01	.05	.08	-.01	-.08
Ovrrater Sub	.21**	.13*	.13*	.00	.07	.13*	.10	.13*	.17**	.18**	.23**	-.13*	.12*	.15**	.15**	.06	.15**	.09	-.09	-.23**	.16**	.13*	-.01	.19**	.13*	.20**
Underrater Sub	-.22**	-.15**	-.22**	-.07	-.16**	-.15**	.04	-.14*	-.11	-.06	-.15**	.06	-.09	-.09	-.05	-.06	-.10	-.16**	.07	.24**	-.14*	-.05	-.01	-.17**	-.04	-.10
Agree Sub	.15**	.12*	.19**	.13*	.20**	.10	.04	.10	.07	.00	.06	.08	.12*	.05	.06	.10	.09	.04	-.08	-.08	.17**	.07	.08	.17**	.08	.05
Ovrrater Other	.20**	.15*	.16**	.00	.17**	.11	.07	.07	.17**	.12*	.17**	.01	.12*	.10	.10	.01	.08	.07	-.14*	-.21**	.13*	.10	-.01	.21**	.07	.13*
Underrater Other	-.22**	-.22**	-.22**	-.14*	-.17**	-.26**	-.07	-.13*	-.13*	-.08	-.15*	-.04	-.18**	-.20**	-.06	-.09	-.16**	-.21**	.06	.22**	-.24**	-.14*	-.09	-.23**	-.12*	-.11
Agree Other	.09	.10	.12*	.14*	.16**	.14*	.01	.01	.03	-.03	.00	.15*	.11	.05	-.01	.03	.04	.03	-.07	-.06	.14*	.06	.07	.12*	.08	.02

Note: \*  $p < .05$ ; \*\*  $p < .01$ ; Do = Dominance, Cs = Capacity for Status, Sy = Sociability, Sp = Social Presence, Sa = Self Acceptance, In = Independence, Em = Empathy, Re = Responsibility, So = Social Conformity, Sc = Self Control, Gi = Good Impression, Cm = Communalness, Wb = Well Being, To = Tolerance, Ac = Achievement via Conformance, Ai = Achievement via Independence, Cf = Conceptual Fluency, Is = Insightfulness, Fx = Flexibility, Sn = Sensitivity, Mp = Managerial Potential, Wo = Work Orientation, Ct = Creative Temperament, Lp = Leadership, Ami = Amicability, Leo = Law Enforcement Orientation; See Table 1 for personality facet definitions.