IMPLICATIONS OF ASSESSOR RATINGS OF FEEDBACK SEEKING AND RECEPTIVITY BEHAVIORS ON 360˚ RATINGS OF WORTHY LEADERSHIP AND PERFORMANCE

by

AARON D. FRIEDMAN

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2015

Acknowledgements

First and foremost I would like to thank my parents, Anita and Scott Friedman for their unconditional and unwavering love, encouragement, and support, through graduate school and all endeavors of my life. I love you and am grateful for you as parents. I would also like to thank my extended family for a lifetime of love with plenty of warm and fun memories of family gatherings. Never a dull moment, on either side of the family!

Second, I would like to thank my advisor, Nicolette Lopez, for her support and encouragement since my first day of graduate school as a Master's student in the I/O program. From my first interview during the application process to my dissertation defense, Nicolette has been a reliable and helpful mentor, both enabling and encouraging my successful completion of graduate school, as well as facilitating my life changing introduction to LWF in 2011. I would also like to thank all of my professors from my time at UTA for sharing their expertise with me and investing in my development. Finally, I would like to thank Maria Levings, my Graduate Advisor for all of her help and friendship in navigating the administrative joys of a graduate program.

Third, I would like to thank my Leadership Worth Following family, who are too numerous to be named, but include the Team and Board, both past and present. Each person affiliated with LWF has contributed uniquely and additively to my success as both a graduate student and a consultant. Additionally, this dissertation project would not have been possible without the support and kindness of the leadership and staff of LWF. Specifically, I would like to thank Dale Thompson, Myranda Grahek, Ryan Phillips, and Cara Fay (Wade), for without their work on *The Search for Worthy Leadership*, much of this would not be possible. Specifically, I would like to thank Cara Fay Wade for her support and astute ability to keep me motivated and moving through the final year of research in school.

June 23, 2015

Abstract

IMPLICATIONS OF ASSESSOR RATINGS OF FEEDBACK SEEKING AND RECEPTIVITY BEHAVIORS ON 360˚ RATINGS OF WORTHY LEADERSHIP AND PERFORMANCE

AARON D. FRIEDMAN, PhD

The University of Texas at Arlington, 2015

Supervising Professor: Nicolette Lopez

A study was conducted to evaluate the role of a feedback-orientation-like construct (Behavioral Feedback Focus) in predicting worthy leadership, performance, and developmental outcomes in a sample of leaders who had undergone executive assessment and coaching at Irving, TX-based Leadership Worth Following, LLC. Results of the study suggest that a construct similar to feedback orientation, Behavioral Feedback Focus, is implicitly measured by the Worthy Leadership Model, and that this construct is both conceptually and empirically related to the Capacity, Commitment, and Character to lead. Assessor ratings of Behavioral Feedback Focus were found to be related to 360˚ outcomes including ratings of Worthy Leadership, past performance, and development, particularly as measured by the inbox simulation. Results of this study are presented, together with a discussion of methodological and practical implications to assessment processes and the broader fields of executive assessment, executive coaching, and Consulting Psychology.

Table of Contents

CHAPTER 1

INTRODUCTION

1.1 Executive Assessment

Executive assessment is a specific type of psychological assessment, occurring within the domain of assessment center techniques, which seeks to evaluate current leaders and leadership candidates on a range of job-related knowledge, skills, and abilities. Specifically, the management and leadership skills that are often evaluated in an assessment center context include: oral and written communication skills, behavioral flexibility, creativity, tolerance of uncertainty, planning and organization skills, and decision making. Assessment centers often utilize a combination of interviews, psychological tests and inventories, and a variety of situational exercises (in-basket, or inbox, leaderless group discussion, subordinate meeting), which seek to approximate various aspects of a leadership position. Trained raters then analyze and rate the participants' performance against a given competency model. Assessment centers have been shown to reveal valid information regarding necessary qualities and attributes for success, overall serving as valid predictors of job (leadership) performance (Riggio, 2000).

One such leadership assessment center is Irving, TX-based Leadership Worth Following, LLC (LWF). LWF assesses leaders across levels (manager-executive), for both selection and development purposes, typically with a subsequent feedback and coaching session(s), as well. LWF assesses leaders against the Worthy Leadership Model (WLM; Thompson, Grahek, Phillips, & Fay, 2008; Appendix B), which consists of three broad constructs (i.e., Capacity, Commitment, Character), 12 factors, 32 dimensions, and 120 behaviors. The Worthy Leadership Model (Thompson et al., 2008) is the competency model against which leadership performance is evaluated at LWF.

1.2 Assessing Feedback Seeking and Receptivity Behaviors within the WLM

Feedback-related behaviors were initially implicated in the current research as a result of a qualitative review (i.e., Friedman Major Area Paper) on one of the 12 factors of the WLM: the

Capacity to Persevere and Adapt (CPA). The CPA is described as being comprised of three latent constructs: energy, adaptability, and humor (Thompson et al., 2008). Thompson et al. (2008) state that "leaders who embody this factor keep going in the face of obstacles, especially when others would wear down and/or give up" (p. 373). One of the behaviors measured during assessment within the CPA relates to receiving bad news and negative feedback without becoming defensive. Conceptually, it may be that processing feedback in a non-defensive manner is a precursor to effectively adapting to and persevering in the face of the aforementioned obstacles encountered in a leadership role, as described by a model constructed as part of a qualitative review on the CPA factor (see Appendix C). Upon closer inspection, feedback-related behaviors are assessed in the WLM outside of the CPA, as well. In total, 10 WLM assessment behaviors related to the seeking of and receptivity to feedback were identified across the WLM, within each of the three broad constructs (Capacity, Commitment, Character; the factor and dimension breakdown can be seen in Appendix D), thus comprising the proposed Worthy Leadership Model - Behavioral Feedback Focus (WLM-BFF; Appendix E) composite scale.

At LWF, depending on the client, up to 90% of the assessments delivered are for development rather than selection purposes (though some clients do engage LWF strictly for selection assessments). Often, those who participate in the selection assessment process return to LWF for developmental coaching if and when they are hired by the client organization. Within a development setting, executive assessment serves as a method for collecting relevant data, on which to then provide feedback and coaching. The ultimate goal is to help a leader or manager more effectively leverage his or her strengths and develop his or her opportunities or performance deficits. Whereas feedback tends to be regarded as information related to the effectiveness of an individual's behavior (Ilgen, Fisher, & Taylor, 1979), coaching "goes beyond" simply providing feedback, with coaches being described as "collaborative problem solvers and caregivers" (London & Smither, 2002, p.87).

1.2.1 Feedback Orientation

One individual difference that can influence the extent to which one is receptive to both coaching and feedback is an individual's feedback orientation (London & Smither, 2002). Feedback orientation is defined as an individual's receptivity to feedback and the extent to which he or she welcomes guidance and coaching. The six dimensions that make up feedback orientation are (a) liking feedback, (b) behavioral propensity to seek feedback, (c) cognitive propensity to process feedback mindfully and deeply, (d) sensitivity to others' view of one's self, (e) a belief in the value of feedback, and (f) feeling accountable to act on the feedback. Simplified (and described as a performance management cycle), feedback orientation can be conceptualized as one's ability to receive, process, and use feedback. In the receiving stage, feedback orientation can influence the initial emotional reactions experienced by the individual. In the processing stage, feedback orientation influences the management of those emotional reactions, and ultimately determines whether feedback is accepted or rejected. Finally, in the use stage, the behavioral consequences are determined (changes or not).

Following the introduction of feedback orientation to the literature, Linderbaum and Levy (2010) developed the Feedback Orientation Scale (FOS; Appendix F). Development of the FOS resulted in four dimensions of feedback orientation: utility (the belief that feedback is useful to achieve goals), accountability (feeling an obligation to react and follow up), social awareness (a tendency to use feedback to be aware of others' views), and finally, feedback self-efficacy (perceived competence to interpret and respond to feedback appropriately). The authors go on to describe the utility of future research evaluating the FOS in an assessment center context as "Understanding an individual's feedback orientation would provide insight into the coachability of this individual or where the individual may need additional support in responding to feedback" (Linderbaum & Levy, 2010, p. 1399).

Following the development of the FOS, Braddy, Sturm, Atwater, Smither, and Fleenor (2013) conducted a study to validate the FOS in a leadership development specific context. The results of their study lend support to the criterion-related validity of the FOS by demonstrating that

feedback orientation predicted leaders' reactions to 360˚ feedback (Braddy et al., 2013). Literature has indicated that accepting feedback can facilitate outcomes such as learning (Ilgen, Fisher, & Taylor, 1979), motivation (Kluger & Denisi, 1996), and organizational performance (London, 2003). Positive relationships have also been demonstrated between feedback orientation and positive affect (r = .24; Linderbaum & Levy, 2010) and emotional intelligence (r = .32), and indirectly (and non-significantly) related to supervisor-rated task performance (r =.16; Dahling, Chaue, & O'Malley, 2012). Braddy et al. (2013) state, however, that to date, feedback orientation has not shown a direct relationship with work-related outcomes as rated by others (contrasted with self-ratings). They conclude their validation study suggesting that, "Future research should continue to investigate the validity of the FOS to ascertain the role that feedback orientation plays in both performance management and leader development" (p. 712).

CHAPTER 2

CURRENT STUDY

2.1 Research Questions

The purpose of this study was to address two specific calls for future research. The first was to address the utility of measuring feedback-seeking and receptivity behaviors in an assessment center context and the second was to examine those behaviors in relation to leadership (work-related) outcomes as rated by others. In addressing these issues, two primary research questions were evaluated:

I. Is a construct similar to feedback orientation implicitly measured by the Worthy Leadership Model (i.e., WLM-BFF)?

II. Does the WLM-BFF predict leadership (work-related) outcomes as rated by others?

2.2 Hypotheses

To empirically evaluate these research questions, the current research tested four main hypotheses in two separate studies. Study 1 sought to evaluate the first hypothesis, proposing convergence between the Worthy Leadership Model Behavioral Feedback Focus (WLM-BFF) scale and the Feedback Orientation Scale (FOS):

I. H_1: WLM-BFF behaviors will significantly converge with, or relate to, the FOS.

These results will describe the relationships between the FOS and WLM-BFF, and the extent to which the results of study 2 can be related to feedback orientation (and the FOS) as a construct and scale.

The 2nd - 4th hypotheses were evaluated as part of study 2, a non-experimental design conducted via analysis of a dataset constructed from archived leadership assessment results. The second hypothesis evaluated the extent to which the WLM-BFF predicted work-related outcomes as rated by others including past performance, the extent to which leadership worth following was demonstrated, and developmental changes in both performance and worthy leadership, across three different work simulations: inbox, team meeting, and leader meeting. So,

while research has investigated (and not found) relationships to work-related outcomes via self-report ratings of feedback orientation, the current study is an investigation of assessment center scores (WLM-BFF) on work-related (leadership) and developmental outcomes. It may be that assessment center ratings are more valid than self-report ratings for a construct such as feedback orientation, as those low in feedback orientation and high in defensiveness might not be willing to rate themselves low on being open to feedback. This problem may be overcome by utilizing other-ratings (contrasted with self-ratings) of feedback-seeking and receptivity, as the current study investigated. This is in line with Oh, Wang, and Mount's (2011) meta-analytic finding that other-ratings on a measure of personality were a better (more valid) predictor of job performance than self-ratings, and demonstrated meaningful incremental validity over self-report ratings in predicting job performance.

II. H_2: The WLM-BFF will significantly predict work-related (leadership) outcomes rated by others.

The third hypothesis evaluated the extent to which the WLM-BFF predicts changes, or development, from time one (T1) to time two (T2) of a performance rating survey (360˚ & Pulse; items presented in Appendix G). This hypothesis may be the most direct evaluation of the proposed concept that the higher the WLM-BFF scores from the assessment, the more likely there will be a performance or demonstrated leadership change from T1 (baseline) ratings on the 360˚ to T2 (Pulse) survey ratings. The underlying theory is that the ultimate goal of feedback (i.e., receive, process, use) is the *use* of that feedback to make behavioral changes. This hypothesis sought to evaluate the ability of the WLM-BFF to predict the likelihood of those changes occurring:

III. H_3: The WLM-BFF will predict changes from the initial (T1) 360˚ survey to the follow-up (T2) Pulse survey.

Finally, the fourth hypothesis evaluated observations (or perceptions) of change as rated by others. This hypothesis evaluated the utility of the WLM-BFF to predict ratings on a (T2) Pulse

survey item which sought to evaluate the extent to which others in the organization have observed behavioral changes in that leader following the assessment process:

IV. H_4: WLM-BFF will predict the (T2) Pulse survey observation of change item rating (Appendix H).

In order to evaluate these hypotheses, the current dissertation included two independent studies: An online survey study to assess the relationship between the FOS and the WLM-BFF and an archival study to assess the utility of the WLM-BFF to predict leadership performance and development.

CHAPTER 3

STUDY 1 METHOD

Study 1 was conducted by collecting responses from an online survey which included the WLM-BFF scale and the FOS. Responses were evaluated in order to investigate the relationship(s) between the two scales.

3.1 Participants

Participants for the first (online survey) study were recruited via Survey Monkey's data collection service, Survey Monkey Panel. The Panel service allowed for a targeted sample of U.S.-based senior managers and executives, roughly equivalent to the leadership level that was utilized in study 2. Senior managers are organizational leaders of people (others report to them) who are higher in an organization's leadership structure than entry-level managers, but below executive level leaders such as vice presidents and C-Level executives (e.g., CEOs, CFOs, etc.). Specific demographics of this sample, including, age, level of leadership, industry served, number of direct reports, and level of education are presented in Table 1 (Tables presented in Appendix A).

In order to evaluate H_1 via the survey study, 128 participants were recruited via Survey Monkey Panel. The purpose of study 1 was to evaluate the relationship between the FOS and BFF. Utilizing G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) an a priori power analysis was conducted to determine that 84 participants would be required to detect a correlation of .30 (a medium effect size; Cohen, 1988) at a power level of .80. As this study is based on conceptual overlap between the two scales, a medium effect size was expected (at minimum).

3.2 Materials

Study 1 required the administration of two individual scales: The new Behavioral Feedback Focus scale (WLM-BFF) and the previously published Feedback Orientation Scale (FOS). Together, the survey consisted of 30 items (20 FOS and 10 WLM-BFF), plus demographics. The full survey can be seen in Appendix I.

3.21 LWF assessor ratings (WLM-BFF)

The first set of items in the online survey consisted of the 10 WLM-BFF items that were selected by subject matter experts based on conceptual overlap with the FOS (described in more depth in study 2). Item stems were taken from the BFF behaviors and converted into a self-report format via the instructions of rating the extent to which the participant demonstrates the behaviors. The individual items were rated on a 5-point Likert-type scale, generating item-level ordinal data, though aggregated to create a continuous variable, for the overall BFF score.

3.22 Feedback orientation scale (FOS)

The second set of items administered via the online survey was the 20-item FOS (Linderbaum & Levy, 2010). The primary purpose of administering this scale was to evaluate convergence between the FOS and the WLM-BFF. The FOS outlines four dimensions ultimately to be measured: utility (α = .86), accountability (α = .74), social awareness (α = .80), and feedback self-efficacy (α = .77). Five self-report items are present within each dimension for a total of 20 items and a previously demonstrated overall alpha of .86. Items are rated on a 5-point scale, anchored at either end with (1) *strongly disagree* to (5) *strongly agree*. Similar to the BFF scores, the ordinal item-level data were aggregated into a total (average) or whole score, thereby generating a distribution of continuous scores, which was then correlated with the WLM-BFF scale.

3.3 Procedures

In order to evaluate the relationship between the new WLM-BFF scale and the FOS, both scales were loaded into Survey Monkey and administered to a targeted sample of 128 U.S.-based senior managers and executives. The survey was expected to take no longer than 10 minutes. Responses were evaluated for inclusion to screen for surveys that were completed too quickly or that have incomplete or clearly erroneous data (straight 3s, for example). Data were downloaded and imported to SPSS for analysis.

3.4 Analyses

To evaluate the relationship between the WLM-BFF scale and the FOS, a bivariate correlation analysis was conducted in SPSS between the overall BFF and FOS scores. Post-hoc analyses were conducted to determine which factors of the FOS (utility, accountability, social awareness, feedback self-efficacy) were most highly related to the WLM-BFF scale and/or to items within each WLM construct (Capacity, Commitment, Character). However, as the BFF and FOS are intended for whole score interpretation, the full-score correlation was the primary consideration.

Additionally, responses to the item-level WLM-BFF scale were evaluated for psychometric properties including internal reliability and item total correlations. Sample adequacy and appropriateness for a principle component (exploratory factor) analysis were evaluated via a Kaiser-Meyer-Olkin Measure of Sampling Adequacy and Bartlett's test of sphericity, in line with procedures described by Field (2005).

CHAPTER 4

STUDY 1 RESULTS

4.1. Convergence

Hypothesis 1 stated that WLM-BFF behaviors would significantly converge with, or relate to, the FOS. A bivariate correlation was conducted between the self-report WLM-BFF scale aggregated score and the FOS aggregated score. Results of the analysis indicated that the scores converge (r = .63, p < .001), indicating a large effect size (Cohen, 1988). Therefore, hypotheses 1 was supported and the null hypothesis was rejected.

4.2 Reliability

An evaluation of the WLM-BFF's properties was conducted, and results of the analysis suggest that the scale was internally consistent (α = .89). Item and item-total statistics are presented in Table 2. Psychometric properties of the FOS were also evaluated for consistency with previous validation studies (Linderbaum & Levy, 2010). Results of the analyses indicated the whole scale was internally consistent (α = .93). Item and item-total statistics of the FOS are presented in Table 3.

4.3 Principal Components

A principal components analysis (PCA) was conducted on the WLM-BFF and the suitability of PCA was assessed prior to analysis. The overall Kaiser-Meyer-Olkin (KMO) measure was .89, indicating that the sample was adequate for dimension reduction (Kaiser, 1974). Bartlett's test of sphericity was statistically significant, χ^2 (45, N = 10) = 615.15, p < .001, indicating that the data were likely factorizable, as well. PCA revealed one component that had an eigenvalue greater than one and which explained 52.67% of the total variance. Visual inspection of the scree plot (Appendix J) indicated that one component should be retained (Cattell, 1966), as well. As such, one component was retained (i.e., Behavioral Feedback Focus).

4.4 Conclusion

Because a significant relationship was found between the whole scores of the WLM-BFF and FOS, post hoc analyses were conducted to evaluate the relationships and internal consistency at the factor level of the FOS and the WLM-BFF, as well. Despite a single component extraction from the WLM-BFF, due to the construction of the composite scale from each of the three competencies of the WLM (Capacity, Commitment, Character), dimension -level aggregates were calculated at the competency level and evaluated in relationship to the whole-scale score, as well as the factor level of the FOS. Results of analyses are presented in Table 4.

CHAPTER 5

STUDY 2 METHOD

Study 2 was the primary study in the current research to evaluate the research question as to whether the WLM-BFF could predict performance, worthy leadership, and changes in performance (development) following participation in executive assessment, including receiving subsequent feedback and coaching based on the results of that assessment.

Study 2 was conducted via an archival dataset constructed from assessment results obtained from Irving-based consulting firm Leadership Worth Following, LLC. Assessor ratings on three individual work simulations were examined as predictors of ratings of worthy leadership and past performance on a T2 360° survey, and predictors of change from baseline across both worthy leadership and performance.

5.1 Participants

Participants in study 2 consisted of 118 senior managers (between entry level managers and executive level leaders) who were in a people-leadership position (i.e., others report to them) and had undergone developmental assessment at Leadership Worth Following, LLC. Participants were recruited through LWF being selected as the assessment vendor by the organization whose leaders' data were used in this study. In addition to the developmental feedback received, at the time of assessment, participants provided consent for LWF to use the data generated during their assessment for research purposes. Participants included in this study were selected from one organization (to ensure consistency in the assessment process and subsequent feedback components) which is a large DFW-based transportation organization. While locally headquartered, the senior managers who participated in the development assessment (and whose data will be used in analysis) were from locations around the country. This sample was also inclusive of senior managers across departments, including operations, marketing, finance, etc.

Inclusion criteria of the proposed study narrowed the available sample significantly. Ultimately, the limiting factor was those participants who had completed each component of the assessment necessary for inclusion in this study. In this particular case, the limiting variable was the (T2) Pulse survey (as some leaders who have been assessed have not yet taken the T2 survey, thus were not be eligible for inclusion in this study), for a total sample size of 118. Demographic information of the sample is presented in Table 5.

5.2 Materials

The measurement tools used in study 2 are as follows: (a) WLM-BFF assessor ratings, (b) 360˚ feedback survey (T1), and (c) Pulse feedback survey (T2).

5.2.1 WLM-BFF assessor ratings.

The LWF assessor ratings that were used in this study were generated as part of LWF's standard assessment process, which includes but is not limited to three work simulations including an inbox, team meeting, and leader meeting. The inbox is an email-based work simulation where participants are evaluated on their ability to work through a dense, online email simulation task in a time- and resource-limited situation. The inbox simulation provides an opportunity for the participant to demonstrate their ability to make decisions, collect new information, set priorities, inform and involve others, and take action on a variety of items and issues identified in email messages, faxes, letters and voice mail messages. The team meeting consists of two assessors participating in a leaderless group discussion with the participant to evaluate their ability to work as a member of a team, manage conflict, and include others' perspectives in a decision-making process. The participant must collect and incorporate new information, lead the group to set priorities, make mutually beneficial decisions, and take action on a variety of items centered on evaluating two alternative business initiatives. Finally, the leader meeting consists of one assessor playing the role of the participant's boss, or leader, where the participant is required to present an overview of a business situation and answer various follow-up questions from that leader. The leader meeting role play simulation provides an opportunity for the participant to demonstrate the ability to effectively present a business case to

a leader. The participant collects and incorporates new information, proposes decisions and priorities, while credibly representing his/herself and the organization. For the inbox simulation, the participant is given two hours to complete the task. For each in-person work simulation (team and leader meetings), the participant is given 30 minutes to prepare, followed by the simulation, which lasts for 30 minutes. Immediately following participation in each of the work simulations, the primary assessor scores the participant's performance against 139 total WLM behaviors, organized by the aforementioned construct-factor-dimension-behavior model (Appendix B). In addition to the 120 standard behaviors in the WLM, assessment ratings at this level and for the client include 19 additional behaviors as part of the assessment. It is from these 139 assessment ratings across each of the four work simulations that the WLM-BFF is extracted. For each participant, each simulation (multi-method) is rated by a different assessor (multi-rater). Ratings are based on a 1 to 5 scale, at half-point intervals, with behavioral anchors at the 1, 3, and 5, describing below average, average, and above average manifestations of the behavior being rated.

In order to select the WLM-BFF specific behaviors for this study, the aforementioned 139 behaviors were evaluated for conceptual overlap with and relation to seeking of and receptivity to feedback. Ten behaviors were ultimately chosen as related to a conceptual "feedback focus." While feedback seeking and receptivity behaviors may manifest differently across the three work simulations, each of the three simulations in the current study provide opportunities for the participant to demonstrate these behaviors. For example, in the inbox, the participant may take the opportunity to seek feedback and check assumptions on an ambiguous project status report prior to submitting it to his or her leader. In the team meeting, the participant may seek, and subsequently use feedback from the other role players regarding the current status of the business and which solution option might provide the highest likelihood of success. Finally, in the leader meeting, the participant is tasked with a presentation, and therefore must proactively seek feedback and commentary on that presentation in order to demonstrate these behaviors. It is interesting to note that of the 10 total behaviors, six of those behaviors fall within the

"Commitment to Lead" construct of the WLM (Thompson et al., 2008). The reason this is noteworthy is because Commitment is described as what leaders "want to do" whereas Capacity is what they "can do" and Character is what they "will do". As what a leader would "want to do" could be considered a type of motivation, this may already demonstrate some additional theoretical convergence with the FOS, as scores on the FOS in previous research were shown to significantly relate to achievement motivation (r =.45, p < .01; Braddy et al., 2013).

5.2.2 360˚ feedback survey (Baseline)

The DVs of the analyses came from two separate 360˚ surveys: the initial 360˚ survey and the follow-up (T2) Pulse survey. The 360˚ survey is administered prior to the assessment process so that results are available at the time of feedback, and is structured so that the participant first completes a performance survey to generate self-report ratings across their appropriate leadership model. The participant then provides the assessment coordinator at LWF a list of their raters to invite to participate in the 360˚ survey. Raters are designated as either a direct report, peer, primary leader, secondary leader, or "other." In total, 76 behaviors are rated on a 5-point scale anchored by how frequently the behavior is demonstrated: *to no extent* (1) *to a very great extent* (5). Ratings for each behavior are presented aggregated by rater group, along with self-ratings, and an overall score which does not include the self-rating.

In addition to the 76 behaviors rated against the organization's leadership model/structure, two additional items are assessed: *How would you rate this individual's overall performance during the past year* (1 to 5 scale, from *below average* to *significantly above average*) and *Please indicate the extent to which this individual demonstrates leadership worth following* (1 to 5 scale, from *to no extent* to *a very great extent*; Appendix G). These items were utilized in the proposed research as outcome measures for two specific reasons: (a) they are independent of the leadership behaviors that appear in both the predictors (assessment behaviors), as well as the 360˚ behaviors being rated, thus the confound of the same behaviors being used as IVs contributing to the DV scores is removed and (b) they are the only two items

that are guaranteed to be rated on both the (T1) 360˚ and (T2) Pulse survey. This will allow for greater integrity and consistency of the data enabling comparisons from T1 to T2 responses.

5.2.3 Pulse 360˚ survey (Time 2)

Approximately 14 months after the initial assessment and administration of the 360˚ survey, participants are invited to participate in a follow-up "Pulse" survey, which can be thought of as a "mini-360˚." The Pulse survey differs from the initial 360˚ mainly in the number of behaviors rated. Rather than assessing the entire competency model as the initial survey does, participants are asked to choose between 5 and 15 behaviors from the broader model on which they are interested in receiving feedback. Once those behaviors are chosen, the survey is generated for self-report as well as for the selected raters to complete. In addition to the behaviors selected, ratings on the two additional items from the initial survey (performance and demonstration of leadership worth following; Appendix G) are solicited. Additionally, a third item inquiring about the extent of observed post-assessment behavioral change is rated: *To what extent have you observed this individual make positive changes in his/her behavior over the last year* (1 to 5 scale, from *to no extent* to *a very great extent*; Appendix H). As a result of the opt-in rating method of this tool, there is no guarantee or consistency to which behaviors will be selected, aside from the two aforementioned items which also appear on the (T1) 360˚. Again, it is, in part, for that reason that these are the items being assessed on each survey, rather than those at the behavioral level.

5.3 Procedure

The hypotheses in study 2 ($H_{2\text{-}4}$) were tested via an archival dataset constructed from assessment results described above. The archival data included assessor ratings on each of the WLM-BFF behaviors across three work simulations (inbox, team meeting, leader meeting) and two work-related outcome ratings as rated by others (direct report, peer, leader) via 360˚ ratings on an initial 360˚ (T1) and follow-up Pulse (T2) survey, as well as a third item that exists only on the Pulse (T2) survey. The dataset was constructed by assigning participants a novel database

ID (DBID) and calculating their total BFF score on each of the three simulations, creating a new continuous variable as a result of aggregating the assessor ratings on each of the 10 items for use as an independent variable in the analyses. Similar to the procedure described above, averages (whole scores) were calculated only with the ratings present. Dependent variables were constructed by aggregating ratings at the rater-level group (direct report, peer, leader) to create one continuous DV score per rater group, per item, matched by participant DBID.

5.4 Analyses

Independent variables were constructed via the aforementioned method of aggregating assessor scores on the 10 BFF behaviors, within three individual work simulations: inbox, team meeting, and leader meeting. The sample allowed for the computation of 118 total BFF scores, across each of the three work simulations including the inbox (M = 3.09, SD = .32, α = .68), team meeting (M = 2.96, SD = .54, α = .92), and leader meeting (M = 2.99, SD = .39, α = .80). Full descriptives of the independent variables can be seen in Table 6. Zero-order correlations between the independent variables can be seen in Table 7.

Control variables were computed from the results of a 360˚ survey tool administered prior to assessment and served as a baseline, or control, in the current study. Survey results from two items (past performance and extent to which leadership worth following was demonstrated) were used in the analyses. Scores were aggregated when more than one rater was present within a group (direct reports and peers), and leader-level ratings were left as individual ratings (as each subject/participant only had one leader). As a result, a total of six control variables were computed and utilized in the analyses. Descriptives of the control variables can be seen in Table 8 and correlations between control variables are presented in Table 9.

Dependent variables were computed from data collected from the Pulse 360˚ survey tools described above, as well as from calculating the difference in scores between the Pulse survey and the control variables. The Pulse survey variables included two individual items related to past performance and demonstration of leadership worth following (Appendix G) across three rater groups (direct reports, peers, leaders) for a total of six performance outcome-related

variables. Descriptives of the performance and leadership variables are presented in Table 10. Additionally, difference scores were calculated between the two aforementioned DVs, which were also rated in the initial (baseline) 360˚, across three rater groups. Descriptives of the difference scores are presented in Table 11. Finally, the (T2) Pulse survey item related to post-assessment observed behavioral change (Appendix H) was utilized and descriptives are presented in Table 12. Correlations between DVs are presented in Table 13.

Analyses were performed using SPSS REGRESSION and SPSS EXPLORE for evaluation of assumptions for standard multiple regression. With a total of 118 participants (though the total *N* fluctuates across models) and four IVs (three BFF IVs + control variable in each model), the number of cases is within range of Green's "rule of thumb" (1991, p.123) that the required sample size can be determined with the following equations: $N \geq 50 + 8m$ (m = number of IVs) for testing multiple correlation (Tabachnick & Fidell, 2007). Variables were examined for normality of distributions based on measures of central tendency and visual inspection of histograms and boxplots produced in the SPSS EXPLORE function. Variables were examined for outliers based on boxplots produced in the SPSS EXPLORE function. While present, it was determined that the outliers were still within the valid range of scores, and were not adjusted. Assumptions of normality, linearity, and homoscedasticity of residuals were evaluated based on visual inspection of scatterplots. Multicollinerarity was evaluated via examination of Variance Inflation Factor and Tolerance values, and it was determined that the variables met the required assumptions to test the hypothesis in standard multiple regression, across the three hypotheses within study 2. To examine the change score dependent variables in hypothesis 3, while taking initial performance into consideration, multiple regression models were used to evaluate the unique variance accounted for by the WLM-BFF scale(s) by utilizing T1 performance as a predictor variable as well, in line with the statistical adjustment process outlined by Tabachnick and Fidell (2007) for contingencies among IVs in multiple regression.

CHAPTER 6

STUDY 2 RESULTS

6.1 Hypothesis 2

H_2 sought to assess the utility of the BFF score to predict ratings of performance and worthy leadership approximately 14 months following assessment. H_2 was tested via standard multiple linear regression with the BFF scores across the three work simulations as the independent variables (IVs) and the two Pulse (T2) items (performance and worthy leadership) across three rater groups (direct report, peer, leader) as the dependent variables (DVs). The independent variables (inbox BFF, team meeting BFF, leader meeting BFF) were tested together (as multiple predictors) in relation to each outcome variable across each of the three rater groups. As hypothesis 2 sought to evaluate ratings 14 months post-assessment, without evaluating a baseline score, testing this hypothesis was an evaluation of the utility of the WLM-BFF scale in a selection situation (different from development purposes), where initial 360˚ scores would not be available to include in a predictive model, but assessor ratings (of BFF) would be.

6.1.1 Performance

With regard to performance at the direct report rater level, Table 14 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variability in the performance rating by direct reports, F (3, 100) = 2.93, p = .07. However, higher assessor ratings on the inbox WLM-BFF scale did predict higher ratings on performance as rated by direct reports. The team meeting and leader meeting BFF scores did not account for additional unique variance in the model, at the direct report rater level (Table 14).

With regard to performance at the peer rater level, Table 15 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95%

confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variability in the performance rating by peers, F (3, 107) = 2.52, p = .06. However, higher assessor ratings on the inbox WLM-BFF scale did predict higher ratings on performance as rated by peers. The team meeting and leader meeting BFF scores did not account for additional unique variance in the model, at the peer rater level (Table 15).

With regard to performance at the leader rater level, Table 16 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variance in performance scores at the leader rater level, F (3, 113) = 1.09, p = .36. Additionally, none of the simulation BFF scores (inbox, team meeting, leader meeting) accounted for additional unique variance in the model at the leader rater level (Table 16).

6.1.2 Leadership worth following (worthy leadership)

With regard to worthy leadership at the direct report rater level, Table 17 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variance in worthy leadership, F (3, 100) = 2.59, p = .06. However, higher assessor ratings on the team meeting WLM-BFF scale did predict higher ratings of worthy leadership at the direct report rater level. The inbox and leader meeting BFF scores did not account for additional unique variance in the model at the direct report rater level (Table 17).

With regard to worthy leadership at the peer rater level, Table 18 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variability in worthy leadership ratings at the peer rater level F (3, 108) = 2.50, p = .06. However, higher assessor ratings on the inbox WLM-BFF scale did predict higher ratings on

worthy leadership as rated by peers. The team meeting and leader meeting BFF scores did not account for additional unique variance in the model at the peer rater level (Table 18).

With regard to worthy leadership at the leader rater level, Table 19 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variance in worthy leadership ratings at the leader rater level, F (3, 114) = 2.07, p = .11. Higher assessor ratings on the inbox WLM-BFF scale did predict higher ratings on worthy leadership as rated by leaders. The team and leader meeting BFF scores did not account for additional unique variance in the model at the leader rater level (Table 19).

As the BFF scores from the inbox accounted for unique variability in performance ratings at the direct report and peer level, as well as for worthy leadership ratings at the peer and leader levels, and BFF scores from the team meeting accounted for unique variability in worthy leadership ratings at the direct report rater level, hypothesis 2 was partially supported. Because higher scores in these simulations were related to higher Pulse ratings, the relationships demonstrated were in the expected direction. Within the context of selection assessment, when baseline 360˚ scores are unavailable, utilizing assessor ratings on the WLM-BFF scale from the inbox and team meeting may contribute to positively predicting certain Pulse survey ratings, at certain rater levels, approximately 14 months after assessment.

6.2 Hypothesis 3

H_3 evaluated the utility of the WLM-BFF to predict changes from the 360˚ (T1) ratings to the Pulse (T2) survey ratings of those same items (performance and worthy leadership). In order to conduct this analysis, change (difference) scores were calculated from the T1 to T2 DVs, creating two new variables across three rater groups. These new variables served as continuous dependent variables, with the BFF scores on each of the three simulations as the independent variables and the relevant baseline (T1) 360˚ scores as the control variables. Whereas hypothesis 2 was an evaluation of a selection situation when baseline 360˚ scores are

unavailable, this hypothesis focused on predicting development, taking baseline ratings of performance and worthy leadership into account, as would be the purpose of a developmental assessment preceding a coaching or development planning situation.

6.2.1 Change in performance

With regard to change in performance at the direct report rater level, Table 20 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 37% of the variability in change in performance at the direct report rater level, F (4, 84) = 14.09, p < .01. After controlling for baseline performance ratings, higher assessor ratings on the inbox WLM-BFF scale predicted more improvement in performance as rated by direct reports. Lower baseline ratings on performance by direct reports also predicted improvement in performance ratings at the direct report rater level. The team meeting and leader meeting BFF scores did not account for additional unique variance in the model at the direct report rater level (Table 20).

With regard to change in performance at the peer rater level, Table 21 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 33% of the variability in changes in performance at the peer rater level, F (4, 102) = 14.21, p < .01. After controlling for baseline performance ratings, higher assessor ratings on the inbox WLM-BFF scale predicted more improvement in performance as rated by peers. Lower baseline ratings on performance by peers also predicted improvement in performance ratings at the peer rater level. BFF scores from the team meeting and leader meeting did not account for additional unique variance in the model at the peer rater level (Table 21).

With regard to change in performance at the leader rater level, Table 22 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations

(sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 37% of the variability in changes in performance at the leader rater level, F (4, 109) = 17.63, p < .01. After controlling for baseline performance ratings, none of the simulation BFF scores (inbox, team meeting, leader meeting) accounted for additional unique variance in the model at the leader rater level; however, lower baseline ratings on performance by leaders did predict improvement in performance ratings at the leader rater level (Table 22).

6.2.2 Change in leadership worth following (worthy leadership)

With regard to change in worthy leadership ratings at the direct report rater level, Table 23 displays the correlations between the variables, the unstandardized regression coefficients (B) and intercept, standard error for B, the standardized regression coefficients (β), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 39% of the variability in change in worthy leadership as rated by direct reports, F (4, 84) = 14.94, p < .01. After controlling for baseline worthy leadership ratings, none of the simulation BFF scores (inbox, team meeting, leader meeting) accounted for additional unique variance in the model at the direct report rater level; however, lower baseline ratings on worthy leadership by direct reports did predict improvement in worthy leadership ratings at the direct report rater level (Table 23).

With regard to change in worthy leadership at the peer rater level, Table 24 displays the correlations between the variables, the unstandardized regression coefficients (B) and intercept, standard error for B, the standardized regression coefficients (β), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 38% of the variability in change in worthy leadership at the peer rater level, F (4, 103) = 17.39, p < .01. After controlling for baseline worthy leadership ratings, assessor ratings on the inbox WLM-BFF scale predicted more improvement in worthy leadership at the peer rater level. Greater improvement was also predicted by lower baseline worthy leadership ratings by peers. BFF scores from the team meeting and leader meeting did not account for additional unique variance in the model at the peer rater level (Table 24).

With regard to change in worthy leadership ratings at the leader rater level, Table 25 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 49% of the variability in changes in worthy leadership at the leader rater level, F (4, 110) = 28.17, $p < .01$. After controlling for baseline worthy leadership ratings, none of the simulation BFF scores (inbox, team meeting, leader meeting) accounted for additional unique variance in the model at the leader rater level; however, lower baseline ratings on worthy leadership by leaders did predict improvement in worthy leadership ratings at the leader rater level. (Table 25).

The inbox BFF score accounted for unique variability in change scores in performance at the direct report and peer levels, as well as worthy leadership at the peer level. Higher inbox BFF scores were associated with greater improvement from baseline whereas lower baseline ratings predicted more improvement in both performance and worthy leadership across the models. As such, hypothesis 3 was partially supported with regard to the utility of the WLM-BFF scale, together with baseline 360˚ ratings, to predict (positive) developmental change as measured by 360˚ ratings.

6.3 Hypothesis 4

H_4 evaluated the utility of the WLM-BFF to predict the Pulse (T2) survey item related to perceptions of post-assessment behavioral change observed. This hypothesis evaluated ratings of perceptions of change, rather than change calculated from differences in baseline ratings, though baseline 360˚ ratings were taken into consideration in the evaluation of this hypothesis to control for baseline worthy leadership and performance.

6.3.1 Observed behavioral change - direct reports

With regard to observed change in behavior at the direct report rater level, Table 26 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not

account for a significant portion of the variability in ratings of observed change by direct reports, F (4, 84) = .37, p = .83. After controlling for baseline performance ratings, neither the simulation BFF scores (inbox, team meeting, leader meeting), nor the baseline performance ratings by direct reports, accounted for additional unique variance in the model at the direct report rater level (Table 26).

6.3.2 Observed behavioral change – peers

With regard to observed change in behavior at the peer rater level, Table 27 displays the correlations between the variables, the unstandardized regression coefficients (B) and intercept, standard error for B, the standardized regression coefficients (β), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variability in ratings of observed change by peers, F (4, 102) = 1.68, p = .16. After controlling for baseline performance ratings, neither the simulation BFF scores (inbox, team meeting, leader meeting), nor the baseline performance ratings, accounted for additional unique variance in the model at the peer rater level (Table 27).

6.3.3 Observed behavioral change – leader

With regard to observed change in behavior at the leader rater level, Table 28 displays the correlations between the variables, the unstandardized regression coefficients (B) and intercept, standard error for B, the standardized regression coefficients (β), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variability in ratings of observed change by leaders, F (4, 105) = .32, p = .86. After controlling for baseline performance ratings, neither the simulation BFF scores (inbox, team meeting, leader meeting), nor the baseline performance ratings, accounted for additional unique variance in the model at the leader rater level (Table 28).

Hypothesis 4 was also evaluated using baseline worthy leadership ratings (as opposed to performance) as a control. No differences in results were detected. The correlations between the variables, unstandardized regression coefficients (B) and intercept, standard error for B, the standardized regression coefficients (β), the semipartial correlations (sr_i^2), 95% confidence

interval, R^2, and adjusted R^2 are presented for direct report (Table 29), peer (Table 30), and leader (Table 31) rater groups. As none of the BFF scores across simulations accounted for unique variability in ratings on the observed behavioral change item, hypothesis 4 was not supported and the null hypothesis was retained.

6.4 Between-Simulation BFF Scale

Post hoc analyses were conducted by constructing a BFF scale at the behavioral level in relation to the 360˚ outcomes, across the three work simulations (as opposed to within). Relationships to outcomes (whether performance, worthy leadership, or developmental change) were evaluated to determine which work simulation produced the greatest frequency of relationships to the outcomes of interest at the behavioral level, thereby identifying the work simulation that is measuring the WLM-BFF behavior of interest most effectively (with the greatest frequency), as it may relate to the outcomes of interest in the current study. Behavioral-level frequencies are presented in Table 32.

Following the identification of which work simulations from which to pull individual behaviors, those behaviors (across the work simulations) were aggregated to produce a new independent variable, the "super IV," or the between-simulation BFF scale. Descriptive statistics of the new variable are presented in Table 33. The super IV was evaluated as a single independent variable (together with a control variable if appropriate), with the outcomes previously examined as dependent variables in a series of 18 regression models, across the three hypotheses tested in study 2.

6.4.1 Hypothesis 2

Performance.

With regard to performance at the direct report rater level, Table 34 displays the unstandardized regression coefficient (*B*) and intercept, standard error for *B*, the standardized regression coefficient (*β*), and 95% confidence interval. The super IV did not account for a

significant portion of the variance in performance scores at the direct report rater level $F(1, 96) = .50$, $p = .48$, $R^2 = .01$.

With regard to performance the peer rater level, Table 35 displays the unstandardized regression coefficient (*B*) and intercept, standard error for *B*, the standardized regression coefficient (*β*), and 95% confidence interval. The super IV did not account for a significant portion of the variance in performance scores at the peer rater level $F(1, 103) = .35$, $p = .55$, $R^2 = .003$.

With regard to performance at the leader rater level, Table 36 displays the unstandardized regression coefficient (*B*) and intercept, standard error for *B*, the standardized regression coefficient (*β*), and 95% confidence interval. The super IV did account for a significant portion of the variance in performance scores at the leader rater level $F(1, 109) = 8.63$, $p = .004$, $R^2 = .07$. Higher assessor ratings on the super IV predicted higher ratings on Pulse survey performance, as rated by leaders.

Leadership Worth Following.

With regard to worthy leadership at the direct report rater level, Table 37 displays the unstandardized regression coefficient (*B*) and intercept, standard error for *B*, the standardized regression coefficient (*β*), and 95% confidence interval. The super IV did not account for a significant portion of the variance in worthy leadership ratings at the direct report rater level $F(1, 96) = .09$, $p = .77$, $R^2 = .001$.

With regard to worthy leadership at the peer rater level, Table 38 displays the unstandardized regression coefficient (*B*) and intercept, standard error for *B*, the standardized regression coefficient (*β*), and 95% confidence interval. The super IV did not account for a significant portion of the variance in worthy leadership ratings at the peer rater level $F(1, 104) = 3.46$, $p = .07$, $R^2 = .03$.

With regard to worthy leadership at the leader rater level, Table 39 displays the unstandardized regression coefficient (*B*) and intercept, standard error for *B*, the standardized regression coefficient (*β*), and 95% confidence interval. The super IV did account for a significant portion of the variance in worthy leadership ratings at the leader rater level $F(1, 110) = 11.16$, p

= .001, R^2 = .09. Higher assessor ratings on the super IV predicted higher worthy leadership ratings on the Pulse survey, as rated by leaders.

As the WLM-BFF super IV significantly predicted performance and worthy leadership ratings at the leader rater level (only), hypothesis 2, when tested with the super IV, was partially supported.

6.4.2 Hypothesis 3

Change in Performance.

With regard to change in performance at the direct report rater level, Table 40 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 33% of the variability in change in performance at the direct report rater level, F (2, 80) = 21.35, p < .001. After controlling for baseline performance ratings, however, higher assessor ratings on the super IV did not predict improvement in performance as rated by direct reports. Lower baseline ratings on performance by direct reports did predict improvement in performance at the direct report rater level.

With regard to change in performance at the peer rater level, Table 41 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 30% of the variability in change in performance at the peer rater level, F (2, 98) = 22.70, p < .001. After controlling for baseline performance ratings, however, higher assessor ratings on the super IV did not predict improvement in performance as rated by peers. Lower baseline ratings on performance by peers did predict improvement in performance ratings at the peer rater level.

With regard to change in performance at the leader rater level, Table 42 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept,

standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 43% of the variability in change in performance at the leader rater level, F (2, 105) = 40.74, $p < .001$. After controlling for baseline performance ratings, higher assessor ratings on the super IV predicted more improvement in performance as rated by leaders. Lower baseline ratings on performance by leaders also predicted improvement in performance ratings at the leader rater level.

Change in Leadership Worth Following.

With regard to change in worthy leadership at the direct report rater level, Table 43 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 40% of the variability in change in worthy leadership at the direct report rater level, F (2, 80) = 28.65, $p < .001$. After controlling for baseline worthy leadership ratings, however, higher assessor ratings on the super IV did not predict improvement in worthy leadership as rated by direct reports. Lower baseline ratings on worthy leadership by direct reports did predict improvement in worthy leadership ratings at the direct report level.

With regard to change in worthy leadership at the peer rater level, Table 44 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 41% of the variability in change in worthy leadership at the peer rater level, F (2, 99) = 36.62, $p < .001$. After controlling for baseline worthy leadership ratings, however, higher assessor ratings on the super IV did not predict improvement in worthy leadership as rated by peers. Lower baseline ratings on worthy leadership by peers did predict improvement in worthy leadership at the peer rater level.

With regard to change in worthy leadership at the leader level, Table 45 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations

(sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model accounted for 54% of the variability in change in worthy leadership at the leader rater level, F (2, 106) = 64.50, p < .001. After controlling for baseline worthy leadership ratings, higher assessor ratings on the super IV predicted more improvement in worthy leadership as rated by leaders. Lower baseline ratings on worthy leadership by leaders also predicted improvement in worthy leadership ratings at the leader rater level.

As the WLM-BFF super IV predicted changes in performance and worthy leadership ratings at the leader rater level (only), hypothesis 3, when tested with the super IV, was partially supported.

6.4.3 Hypothesis 4

Observed Behavioral Change - Direct Reports.

With regard to observed behavioral change at the direct report rater level, Table 46 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variance in observed behavioral change at the direct report rater level, F (2, 80) = .32, p = .73. After controlling for baseline performance, neither the super IV, nor the baseline performance ratings, accounted for additional unique variance in the model, at the direct report rater level.

Observed Behavioral Change - Peers.

With regard to observed behavioral change at the peer rater level, Table 47 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variance in observed behavioral change at the peer rater level, F (2, 98) = 2.46, p = .09. After controlling for baseline performance, neither the super IV, nor the baseline

performance ratings, accounted for additional unique variance in the model, at the peer rater level.

Observed Behavioral Change - Leader.

With regard to observed behavioral change at the leader rater level, Table 48 displays the correlations between the variables, the unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2. The overall model did not account for a significant portion of the variance in observed behavioral change at the leader rater level, F (2, 101) = .66, p = .52. After controlling for baseline performance, neither the super IV, nor the baseline performance ratings, accounted for additional unique variance in the model, at the leader rater level.

Hypothesis 4 was also evaluated using baseline worthy leadership ratings (as opposed to performance) as a control within the context of the super IV. No differences in results were detected. The correlations between the variables, unstandardized regression coefficients (*B*) and intercept, standard error for *B*, the standardized regression coefficients (*β*), the semipartial correlations (sr_i^2), 95% confidence interval, R^2, and adjusted R^2 are presented for direct report (Table 49), peer (Table 50), and leader (Table 51) rater groups. As none of the super IV BFF scores, nor baseline ratings, accounted for unique variability in ratings on the observed behavioral change item, hypothesis 4 was not supported and the null hypothesis was retained.

CHAPTER 7

DISCUSSION

The current research sought to address two primary research questions: Is a construct similar to feedback orientation implicitly measured by the Worthy Leadership Model? If so, does that construct serve as a unique predictor of work-related outcomes? Results of the aforementioned analyses suggest the answer to both questions is, to an extent, yes.

The results of study 1, which evaluated the relationship between the FOS and the WLM-WLM-BFF scale, supported hypothesis 1, the notion that the two constructs are significantly related to one another. These results open the door for a deeper investigation into the construct validity of the WLM-BFF. While a correlation of .63 between WLM-BFF and the FOS is a "large" effect (Cohen, 1988) and supports convergent validity between the two scales, the relationship is not so strong as to claim that the constructs are the same. However, the correlation is above the recommended cutoff for use as a proxy measure (of feedback orientation) in organizational research (r = .50; Carlson & Herdman, 2012). While both scales are intended for whole-score interpretation, evaluation of the four dimensions of the FOS and the three broad constructs of the Worthy Leadership Model may provide insight into where the two constructs are similar, and where and how they may differ (Table 4).

Linderbaum and Levy's (2010) FOS was ultimately reduced, or "streamlined" (p. 1392), to four dimensions: utility, accountability, social awareness, and feedback self-efficacy, and each of the four dimensions was shown to predict unique variance over and above one another. In practice, however, the primary researcher has found more utility in referring to feedback orientation as it is described by London and Smither (2002) as a longitudinal performance management process, which can be simplified, to an extent, as an individual difference in one's ability to effectively receive, process, and use feedback (though also influenced by the organization's feedback culture or environment). Taking a step back from the FOS to the concept of feedback orientation as a whole, considering one's ability to receive, process, and use

feedback again "theoretically converges" with the Worthy Leadership Model's three broad constructs – The Capacity, Commitment, and Character to lead – corresponding to what a leader can do (Capacity), wants to do (Commitment), and will do (Character; Thompson et al., 2008).

The current WLM-BFF scale spans each of the three broad constructs of the Worthy Leadership Model, though the total score was found to most significantly converge with Commitment (r = .96, p < .001), which is in fact, the strongest relationship among any of the dimensions of either the WLM-BFF or FOS (Table 4). While the size of this relationship can be explained by a majority (60%) of the BFF items stemming from the Commitment construct of the WLM (Appendix B), the Commitment "subscale" of the BFF is also the most highly related to the FOS whole scale score (r = .61, p < .001), as well as to each individual dimension of the FOS, compared to the Capacity and Character items (Table 4). And, as previously discussed, as what a leader would "want to do" could be considered a type of motivation, this may also demonstrate some additional theoretical convergence with the FOS, as scores on the FOS in previous research were shown to significantly relate to achievement motivation (r = .45, p < .01; Braddy et al., 2013).

So, from both an empirical and conceptual standpoint, the BFF items from within the Commitment construct of the WLM seem to do most of the heavy lifting when it comes to the scale functioning as a proxy of the FOS. However, it is the opinion of the researcher that disregarding the function and utility of the Capacity and Character items negates, or works against, the essence, and the wisdom, of the Worthy Leadership Model's intent, and the relationship to the broader concept of feedback orientation (outside of the scope of its scale). As London and Smither (2002) describe feedback orientation, in part, as one's ability to receive, process, and use feedback, those three elements map neatly onto the Worthy Leadership Model's framework of Capacity, Commitment, and Character composing three unique, yet instrumental components of effective leadership, namely, leadership that is truly *worth following* (it is not just a model/company name!). It may be that one's ability to receive feedback is related to their commitment, or motivation, to seek feedback and develop oneself through learning. This

would align with four of the six BFF behaviors within Commitment falling under the factor of Commitment to Learning and Personal Growth, particularly the Self-Awareness and Development dimension of the Worthy Leadership Model (Appendix B). One's ability to process feedback mindfully, however, may be more related to the Capacity to Lead, particularly as it relates to the Capacity to Persevere and Adapt factor, specifically the BFF behavior of not reacting defensively, arguably the seed from which this entire program of research grew. Finally, the Worthy Leadership Model cannot be discussed without incorporating character, initially implicated in executive assessment and the field of consulting psychology by the publication of the *Search for Worthy Leadership* (Thompson et al., 2008). While traditional executive assessment and broader predictors of leadership performance (Judge 2002, 2004) have been rooted in more of the capacity and commitment-based predictors, Thompson et al. (2008) suggested that simply the ability and motivation to perform and lead was not enough to guarantee success. In fact, in some cases, those leaders who had demonstrated some of the strongest "traditional" predictors of leadership performance were in fact those who experienced the most catastrophic failures – both at a personal and an organizational level. Further, each and every one of these failures could be traced back to a character-based deviation from the path to success. The way this may relate to the behavioral feedback focus construct, and the broader feedback orientation construct, is that it is possible that feedback can be received and processed, though in the end disregarded, or not used. Again, the Character construct of the Worthy Leadership Model is ultimately predictive of what a leader *will do* and may be the element of the BFF which captures the ultimate intention to "use" feedback to "guide behavior change and performance improvement" (London & Smither, 2002, pp. 81). A saying in executive assessment goes: "What do toddlers, puppies, and executives have in common? They're all eventually going to do whatever it is they want!" Retaining the BFF items within the Character to Lead may be the unique contribution seeking to measure and predict the extent to which the feedback encountered in a leadership role may, or may not, ultimately be used.

So, while the results of study 1 provide empirical support for the "conceptual overlap" between the WLM-BFF and the FOS, the relationship between the BFF and feedback orientation (not just the scale) may be far more complex, and engrained, than simply assigning it an *r* value of .63. And, like the rest of the work consultants/practitioners undertake in consulting psychology, while guided by empirical data and "science," it is the assessor who must make the final determination and interpretation of the applicability of data to a given person or issue. Thus, while the quantitative relationship between the two scales does provide some insight into the relationship between the two concepts, it is only when taking the broader ideas of both worthy leadership and feedback orientation into consideration that one can begin to integrate their various components into a conceptual convergence, of a far greater "effect size" than .63.

The second research question addressed in the current study is whether or not the WLM-BFF scale could, or would, function in a predictive manner over and above the self-report FOS which had been previously validated from a construct and content perspective (Braddy et al., 2013), though without finding any direct relation to performance outcomes in a leadership development context (and thus lacking when it came to predictive validity). Linderbaum and Levy (2010) also suggested that researchers evaluate the FOS "as part of an assessment center to determine how open an individual is to developmental feedback," as "Understanding an individual's feedback orientation would provide insight into the coachability of this individual or where the individual may need additional support in responding to feedback." (p. 1399). Further, Braddy et al. (2013) stated that across three individual studies conducted, "feedback orientation has not yet shown a direct relationship with work-related outcomes as rated by others" (p. 709). Study 2 of the current research sought to address both Linderbaum and Levy's suggestion and to evaluate the utility of the WLM-BFF scale to do what the FOS has not been able to – predict work-related (worthy leadership and performance; H_2) and developmental outcomes (post assessment changes; H_3).

The current research evaluated the utility of the WLM-BFF rated within three individual work simulations to predict both work-related outcomes as rated by others (H_2; such might be the

method utilized in a selection context) as well as changes in both performance and demonstrated leadership worth following from a baseline rating to a T2 Pulse survey approximately 14 months after assessment feedback and coaching had occurred (H_3; such might be the method or equation that would likely be of value within a developmental assessment context). While significant effects were found in relation to both performance and worthy leadership outcomes (H_2) as well as developmental change (H_3), results were, at best, inconsistent across outcomes, simulations, and rater groups. This warrants further discussion and prompts a number of related questions regarding (a) differences across items, (b) differences across rater groups, and (c) differences across simulations from which the WLM-BFF was computed.

7.1 Differences Across Items

In order to address differences across items, the first question to consider is what is the difference in performance and worthy leadership? With the two items consistently and significantly correlated within rater groups (Table 13), what is the difference between the two constructs empirically (as well as conceptually)? Does the Worthy Leadership Model seek to predict performance or worthy leadership if, in fact, they are not one and the same? In one sense, performance and leadership can be thought of as interdependent covariates of one another. High correlations within rater groups is what would be expected, as orthogonal relationships would seem to describe a situation, or person, as being able to perform (in a leadership position) without leading others, or leading others effectively without contributing to or driving business results (performance). This is not to say, however, that encountering such individuals in the assessment process does not occur. Typically, however, these are one-off cases and, in the experience of the researcher, are more often strong individual contributors (task-related performance), but not in a position to [need to] effectively *lead* individuals. An example may be a chief engineer or scientist in a high-tech firm or company who is hired due to his or her deep expertise in one field, rather than raw ability to effectively lead and inspire performance in others. The flip side of the coin, however, is the truly transformational leader, who, while can and does generate buy in and commitment from their followers, often times leads them to either non-results or undesired ones.

To evaluate the difference in performance and leadership one perspective can be examined within the context of adaptability in leadership, again from the same WLM factor (The Capacity to Persevere and Adapt) which opened the door to the current research. Yukl and Mahsud (2010) state that many indicators of flexible or adaptable behavior exist, though one of the better indicators is the extent to which leaders vary their behavior appropriately towards confronting both tasks and people. Having to balance effectively leading others (people), while simultaneously competently producing work (tasks), requires a certain degree of flexibility and adaptability (process). The researchers categorically outline situations in which adaptability may be particularly vital, including transitions to different leadership positions, managing immediate crises, adapting to emerging threats or opportunities, balancing competing values, and stakeholder conflicts. They also outline particular traits that may influence flexibility and adaptability in leadership including cognitive complexity and ability, social intelligence, openness to learning (potentially related to feedback orientation), and leader self-awareness (one of the four dimensions of the FOS; Linderbaum & Levy, 2010; Yukl & Mahsud, 2010).

So, with regard to a quantitative model, the question then becomes, is performance a function of leadership or is leadership a function of performance? In theory, the best fit model may be an embedded one, in which performance is a function of leadership (i.e., Worthy Leadership) – which is itself of a function of weighted values of Capacity, Commitment, and Character. Such a model would not only account for the strong, but not perfect, relationships between performance and leadership (within rater groups), but also the strong (but not perfect) relationships between Capacity, Commitment, and Character items within the WLM-BFF scale, speaking not only to the reliability across constructs, but also back to the importance and the "essence" of the Worthy Leadership Model to ensure leadership is assessed across each of the three constructs. Just as the validation study of the FOS (Linderbaum & Levy, 2010) found that each dimension of the scale accounted for unique variance over and above each of the other dimensions, so might each of the three Cs in the Worthy Leadership Model, within the model more broadly (or within the context of a multi-dimensional/composite subscale). Two more

questions now become relevant (and perhaps areas for future research): If performance is a function of Worthy Leadership, and Worthy Leadership is a function of weighted values of Capacity, Commitment, and Character (a) what then are those β weights in the model, generally and (b) what are they in relation to behavioral feedback focus? While the current research sought to address the latter question, the former, if even possible to quantitatively model, in the opinion of the primary researcher, would be the "holy grail" of executive assessment, and perhaps even consulting psychology more generally (but to the criterion question introduced later in this dissertation…what is the universal, standardized dependent variable against which to validate that equation?).

So, returning to the question of what is the difference between performance and leadership worth following, one may view performance as the business results obtained by the leader. The extent to which they have demonstrated leadership worth following, however, may be the manner through which they obtained those results, or the how, and the answer to which would likely be some [weighted] combination of what the leader can do (Capacity), wants to do (Commitment), and will do (Character).

One outstanding question that remains even after disambiguating performance from leadership is with regard to the Pulse (T2) survey item inquiring about the extent to which positive behavioral change has been observed in the past year (Appendix H). Would post-assessment (and feedback and coaching) change be more indicative of changes in performance or in leadership worth following? Results of this item were null across rater groups (H_4); however, change was detected, and predicted, when using difference scores from T1 to T2 360˚ surveys (H_3) – which forces an evaluation of the contrast between perception(s) and reality(ies) of developmental change. Just as the frog does not leap from the gradually boiling pot of water, perhaps developmental changes are subtle enough that when others seek to rate behavioral change as an outcome, the results are null. However, when changes are calculated from criterion-oriented outcome ratings (performance and worthy leadership, in this case), changes are not only detectable but predictable, as well. This may suggest that the better measure of change

in executive assessment and consulting psychology is "actual" calculated change rather than ratings or perceptions of change, in that the "reality" of change may exist even when perceptions of it do not. Further, results of the analysis in H_3 suggested that the T1, or control score, was the single best predictor of change across rater groups and outcomes (both performance and worthy leadership) though in a negative direction. While those who initially scored lower on the baseline ratings had the most room to positively change or improve, one area for future research could be an exploration of whether those who changed the most (low to high) from baseline were perceived any differently than those who experienced little or no change from baseline. If differences are indeed found between those two groups, one could examine what might determine, and at what level, whether those leaders are "given credit" for making those changes, or developing. In other words, what needs to happen for the frog to not only notice that the water temperature is rising and leap from the pot, but do so in a way that will be noticed by others?

The next question to explore with regard to differences is not within the context of item-level differences, but that of rater-group level differences. If ratings of performance can be used as a predictive criterion (and changes in those ratings), at which level(s) of a leader's team might those ratings be the most useful?

7.2 Differences Across Rater Groups

Evaluating differences in findings across the three most common rater groups in 360˚ feedback systems (Fitzgibbons, 2004) – direct reports (subordinates), peers, and supervisor (leader) – poses the question: What is the driving force behind differences in perceptions of both performance and demonstrations of leadership between the three groups? While the initial solution is that this can be accounted for through the "different samples hypothesis" (Moriarty, 2003, p. 991), previous research has not supported the assumption that different rater groups observe different "samples" of leadership and performance behaviors, though multiple rater groups have been found to provide incremental validity over supervisor ratings. If the idea that different levels of one's team are seeing different "versions" of the leader, rater error (or lack thereof) may also be helpful in explaining the findings of the current research.

It may be that peers provide the "truest" ratings, as they have the least amount of vested interest in either halo or horn type effects affecting their responses and ratings. While competitiveness for promotions or competence could potentially fuel bias, by and large, the rater group that should conceptually be least susceptible to rater error effects would be one's peers, who may also have the best understanding of a given individual's job function or role. Leaders and direct reports, however, may be more susceptible to halo and horn effects, essentially seeking to either "anoint or condemn," as leaders tend to not rate as harshly as other groups (lowest rating on performance in the Pulse being a 3; Table 10), and may even seek to avoid the conflict of providing negative feedback (perhaps a flip side of the coin of feedback orientation being willing to deliver that negative feedback; likely accounted for in both the openness and courage dimensions of the WLM). It is due to this tendency at the leader level that an assessment "pre-call" is part of standard protocol at LWF, to get a sense of what might "really be going on," over and above the (less than always meaningful ratings) that can be provided by leaders. This is consistent with previous findings that multiple rater groups (peers and direct reports) provide incremental validity over supervisor ratings (Moriarty, 2003) within the context of 360˚ survey ratings. A direct relationship was found, however, with regard to worthy leadership at the leader rater level (H_2) when baseline worthy leadership ratings were not controlled for. Based on the proposed relationship between worthy leadership and performance discussed above, this may suggest that the WLM-BFF scale (measured in the inbox) can predict a leader's perceptions of the "how" of performance rather than just the "what" of the results obtained.

The most outstanding question, however, as it relates to the proposed "embedded function" of leadership and performance, is why the WLM-BFF was found to be predictive of performance as rated by direct reports, but not leadership worth following [the inverse of the leader findings above, though it is worth noting that the relationship with the inbox WLM-BFF score and leadership ratings (and change) at the direct report level were approaching significance; see Tables 17 and 23]. So the difference in findings at that level may be as simple as statistical error, as the relationship between the two variables at the direct report rater level is *r*

= .82, $p < .001$. If the difference is meaningful, however, this may suggest that BFF is a better predictor of task performance than of the "how" that task is performed, at the direct report rater level. Could it be that seeking feedback from one's direct reports enables more effective performance, incrementally (even slightly) over perceptions of performing in a way that is "worth following?" To explore this question even further, the final "piece" of the results discussion would be to evaluate differences across work simulations implicated in the current research.

7.3 Differences Across Simulations

The most consistent finding across the current study was that of the three simulations from which the BFF IVs were constructed, the inbox was the most predictive of performance and leadership, and changes in performance and leadership. While the team meeting BFF's score was predictive of ratings of worthy leadership at the direct report level (H_2), which may be related to collaborative and conflict management behaviors demonstrated in that simulation, in a sense, the inbox was the "hero" of the study. The BFF score computed from the inbox simulation accounted for unique variance in a total of seven individual regression models tested (direct report ratings of performance and change in performance, peer ratings of worthy leadership, performance, changes in worthy leadership and performance, and leader ratings of worthy leadership).

At first glance, this may seem counterintuitive, as feedback-seeking and receptivity behaviors sound like social behaviors that may be better assessed in an in-person simulation rather than a written one such as the inbox, especially given the previous finding linking feedback orientation to emotional intelligence ($r = .32$, $p < .001$; Dahling et al., 2010). However, the inbox also provides for the most consistent stimulus across the work simulations implicated in this research, as there is no "error" in terms of how the consultant or role player provides the stimuli to the participant during a live role play. In order to seek feedback in the inbox, the participant must proactively demonstrate that behavior, as opposed to respond and react to cues, intentional or not, from a role player during a live simulation. In addition to consistency in stimuli, the inbox may

have functioned as the best predictor, as it may be the simulation most highly related to cognitive ability, as well.

Two different meta-analyses conducted in 2008 (Roth, Bobko, McFarland, & Buster; Whetzel, McDaniel, & Nguyen) examined adverse impact of selection assessments. Results were consistent in that simulations saturated with cognitive ability (such as an inbox) produce larger subgroup (in this case, race) effects than do less cognitively-based, and more socially-oriented (oral communication and interpersonal skills; Roth et al., 2008) work simulations (such as the team and leader meeting in the current research). Previous internal research at LWF has also supported the notion that the inbox is not only the most highly correlated with cognitive ability, but also most predictive in terms of performance. It is for that reason that often times in shortened or streamlined assessment process at LWF (often times between a test battery assessment and full assessment), in addition to cognitive and personality testing (and a structured interview), the inbox (only) will be administered as the sole work simulation. For that reason, it may have been that cognitive ability, particularly critical thinking ability (CTA; as leadership populations tend to have higher levels of CTA than the general population) was "pulling up" scores across the inbox ratings, including those related to BFF, while those same behaviors may have been less susceptible to "cognitive halo" in the team and leader meetings. As such, Boyce, Corbet, and Adler (2013) state that in order for research such as the current study to be "maximally useful in practice, we ultimately need to determine which constructs are best measured by which methods and at which level of fidelity" (p. 24). Roth et al. (2008) conclude similarly that "constructs appear to matter a great deal and decision makers should consider this when designing selection systems or developing exam components" (p. 655). So then, what might be an even more effective way to build composite scales within LWF and other assessment centers? One idea is to create a "super IV" or composite scale across every component of the assessment, as it relates to the outcome(s) of interest.

CHAPTER 8

FUTURE RESEARCH AND PRACTICE IMPLICATIONS

8.1 Super Independent Variable

The current research initially sought to evaluate the utility of an independent variable computed within three work simulations, not between them. So, while the BFF score from each work simulation was included in each regression model, the composite scores were still within-simulation. One practice-based implication of the concerns expressed by Roth et al. (2008) and Boyce et al. (2013) is that rather than create identical composite scales within simulations that could be vulnerable to saturation effects (which may have been the case with cognitive ability in the inbox in the current study), one way to address this problem, and control for saturation effects, is to build a composite between simulations, and even between simulations and testing (implicating cognitive and personality variables as well). By creating a "master" correlation table at the behavioral level (rather than the whole score composite level), practitioners may be able to identify or explore, in this case, which BFF behaviors from which individual simulations are most predictive of, or related to, the outcomes of interest. By measuring behaviors across Capacity, Commitment, and Character, as well as across the "four sets of eyes" (eyes of the scientist such as cognitive and personality tests, eyes of the assessors such as the work simulations, eyes of the participant such as the structured interview, and eyes of the organization such as a 360˚ survey, which were also the outcome variables in the current study), practitioners may be able to account for even more variance in outcomes of interest, while simultaneously controlling for aforementioned saturation effects. Such was the purpose of the post hoc analyses evaluating the between-simulation BFF scale, in addition to the within-simulation scores initially proposed.

8.1.1 Implications of a between-simulation super IV

The construction of an additional BFF scale across the role plays (rather than within) enabled the prediction of both performance and worthy leadership, as well as development (change) of both outcomes, at the leader rater level. This is significant in that by and large, the

within-simulation BFF scales (inbox, team meeting, leader meeting) were unable to account for a significant portion of variability in ratings at the leader level.

To begin evaluating these results, the preliminary finding of note was the frequencies of relationships to outcomes determined at the behavioral level of the BFF scale (Table 32). It is noteworthy that the two Capacity items of the BFF scale had the highest number of "hits" with outcomes within the inbox simulation, the simulation which may be most saturated with (the measurement of) cognitive ability, as previously discussed. The even 50/50 split between the leader meeting and inbox of the six Commitment items may align with the necessity of (and motivation to) proactively seek and elicit feedback within those simulations, in order to effectively integrate information and execute on the various tasks necessary for effective performance. Largely, the inbox and leader meeting are the two work simulations that are focused on information out - in other words, effective performance in these simulation is dependent on the outflux of information (from the participant to the assessor). Therefore, in order to both *seek* and *receive* feedback in those simulations, it must be *proactively* solicited, in order for it to ultimately be *used*. Finally, as the two BFF items within the Character construct of the WLM were found to be "best" (highest frequencies of relationships to outcomes) in the team meeting, this elicits an evaluation of how Character might be related to effective teamwork. While the two behaviors within Character span both the Personal Integrity and Ethics and Humility, Gratitude, and Forgiveness factors of the WLM, both behaviors are rooted, to an extent, in the concept of humility. Future research may evaluate the extent to which humility, specifically, plays a role in individual performance in a team setting, within the context of being willing, and able, to receive, process, and use feedback.

With regard to the utility of the super IV aggregate score as a performance and developmental predictor, this approach, or construction of the BFF scale, created a variable that effectively predicted outcomes at the leader level (where the other BFF scales did not), but failed to predict scores at the direct report and peer levels (where the other scales did). So, even within the limited variability of leader ratings discussed above, the "across" (contrasted with "down")

approach to composite scale construction may be able to account for variance that is lost in a "down" (within-simulation) methodology, though arguably while failing to capture meaningful variance that otherwise would be in a "down" (or within simulation) process. Perhaps the takeaway from this finding, more so than the super IV composite accounting for variance in leader ratings, is the importance of evaluating assessment results in both a down (within simulation/tool) method, as well as an across (between simulation/tools) methodology, as both approaches to evaluating the data provide unique contributions to the overall assessment result.

8.2 Implications of Measuring Behavioral Feedback Focus (BFF) as a Performance and Developmental Predictor

The broader implications of the ability to measure behavioral feedback focus, and use that construct to predict performance (H_2), such might be the case in a selection assessment, and changes in performance (H_3) within the context of a developmental assessment, align similarly to those discussed by Linderbaum and Levy (2010) who state that measuring feedback orientation at the time of assessment could help coaches and assessors understand how open an individual will be to subsequent developmental interventions (coaching). Should the WLM-BFF scale be utilized to predict coachability in a development context, the implications could be twofold. First, such a measure could help determine "ROI" at an organizational level for assigning a particular leader to a development process (or hiring them into a role that would require on-going development). Second, knowing one's "baseline" of BFF could help guide an on-going coaching relationship post-assessment. Knowing one's individual level of BFF, or feedback orientation, could help the coach begin the relationship by "coaching" on openness to feedback before going into more performance or task-based topics (that the leader may be less receptive to initially). This does implicate a distinction, however, between "potential to grow" and coachability. Potential to grow can be more related to general cognitive ability (in an assessment context as measured by the Watson Glaser Critical Thinking Appraisal, though Critical Thinking while related to general intelligence has been differentiated as an independent and single-factor construct; Watson &

Glaser, 1980). It can also be related to more "learning-based" predictors such as practice and quality of feedback (Gluck, Mercado, & Myers, 2011). So, while an individual's potential to grow (learn and develop) may be dependent on more crystalized traits such as critical thinking ability (i.e., discrimination among inferences, recognition of assumptions, deduction of conclusions, interpretation of evidence to reach accurate conclusion and evaluating the strength of arguments), coachability, or level of behavioral feedback focus, may be a moderator of the relationship between critical thinking and learning or performance outcomes.

Further, by broadening the scope of the operational definition of feedback (outside of the leadership/job performance domain), investigating the proposed hypotheses may be an opportunity to explore and further align best practices and research findings of the larger human learning field, especially around the role of feedback, or knowledge of results, as it relates to practice and performance. Gluck et al. (2011) suggest that developing a specific skill (such as leadership) requires far more than just repetitive practice, and that without consistent, quality feedback, the ability to continue developing a particular skill tends to level off (or never begin). While perhaps intuitive, this concept has been tested experimentally, dating as far back as 1927. Thorndike's (1927) Law of Effect experiment sought to illustrate the importance of feedback by asking participants to draw a line that was exactly three inches long. Not surprisingly, the only participants who were able to improve the accuracy of the length of their lines over time were the ones who received feedback as they got closer to the desired length. In essence, practice does not make perfect without high-quality, accurate, consistent feedback (from a reliable source). Gluck et al. suggest that it is the role of coaches to provide learners with this feedback, or knowledge of results, which can in turn significantly affect the extent to which an individual's skills improve.

Further, beyond individual-level development or growth-based predictors, the ability to measure behavioral feedback focus, and use that construct as a predictor of work-related outcomes (H_2), may have implications on the team and group level, as well. Should leaders be assessed poorly on the WLM-BFF subscale, perhaps in addition to predicting "coachability"

concerns (H_3), the scale could help to focus leaders and their coaches on developing leaders' feedback orientation, both for their own development and that of their teams'. For example, Lawrence and Wiswell (1993) found that after a team-level feedback training intervention with managers, their teams tended to be more focused, team-oriented, and collegial with one another. Further, the managers who practiced giving and receiving feedback reported viewing themselves and their subordinate work groups as more proactive, engaged, and cohesive as a team. This would suggest that, to an extent, feedback-related behaviors could be trainable. Therefore, a WLM-BFF subscale could serve as an indicator to a coach that developing their client's orientation, or focus on receiving (and giving) feedback, may be the first step towards improving additional individual and team-level outcomes. However, this also opens the door to yet another debate in the literature – whether feedback orientation (and its related BFF construct) is an environmentally-primed state or a more deeply engrained, or inherent, crystalized trait.

8.3 State vs. Trait

London and Smither (2002) initially proposed that feedback is part of a longitudinal performance management process that is both influenced by and contributes to not only the individual difference of feedback orientation but the organization's feedback culture, as well. Feedback culture refers to "the organization's support for feedback, including nonthreatening, behaviorally-focused feedback, coaching to help interpret and use feedback, and a strong link between performance improvement and valued outcomes" (p. 81). Braddy et al. (2013) state, however, that little is still known regarding the extent to which feedback orientation is shaped by environmental factors. They state that perhaps, "one's feedback orientation may be enhanced by working in an environment where seeking and acting on feedback are viewed as indicators of leadership effectiveness, whereas it may be diminished when working in an environment where seeking feedback is seen as a sign of weakness or where psychological safety is low" (p.712). Should this be the case, they argue, feedback orientation could be viewed as more of a state than an enduring trait. Linderbaum and Levy's (2010) perspective is that feedback orientation is an individual difference that is generally stable, though can be influenced "to some extent" over time

with efforts (such as coaching) or an environment that "specifically targets changing it" (p. 1399). This is consistent with Lawrence and Wiswell's (1993) results that suggest increasing various behavioral feedback seeking-related behaviors is trainable. It may be that feedback orientation is a trait-like construct that is moderated (strengthened or weakened) significantly by one's environment. From a personality standpoint, feedback orientation, or BFF, may be closely aligned with Tolerance, as assessed by the California Psychological Inventory (Gough & Bradley, 2005). Tolerance is defined as a scale which assesses "attitudes of tolerance, forbearance, and respect for others, stemming from ethical convictions about the worth of all people" (p.6). The high end of the scale, however, is anchored by describing characteristics such as being open-minded and not biased or dogmatic, both of which may describe aspects of both feedback orientation and behavioral feedback focus. From a leadership-specific standpoint, the CPI Coaching Report for Leaders (Manoogian, 2006) identifies a leadership characteristic known as "Capacity for Collaboration." A leader's Capacity for Collaboration score is computed by aggregating Tolerance with Creative Temperament (which identifies "people of an imaginative, creative temperament, with both the need and potential for visualizing new and different ways of doing things"; p.3). The Capacity to Collaborate from a personality-based standpoint (as measured by the CPI 260) is defined as being "open and receptive to the ideas and input of others, as well as to new and different ideas, information sources, and points of view" (p.6). This description seems to capture the essence of both feedback orientation and behavioral feedback focus. Future research, in addition to implicating critical thinking ability, may be able to associate BFF with Capacity to Collaborate, and its individual components (Tolerance and Creative Temperament). Future research could examine the relationships between these personality traits, as well as cognitive/critical thinking ability, and behavioral feedback focus to then empirically determine the extent to which feedback orientation or BFF may be an inherent trait.

CHAPTER 9

LIMITATIONS

The aforementioned "criterion question" is one often encountered in consulting psychology. What is a true criterion, or outcome, that could be measured to validate executive assessment results? Is it promotion, pay, performance evaluations, revenue generated, or something else? To date, there is no one correct answer or gold standard, as meaningful outcomes likely vary by organization (and leader). While 360˚ scores can evaluate, to an extent, "overall leadership performance," the data that are generated by 360˚ tools are not without limitations. For example, the 360˚ data were obtained from peers, direct reports, and leaders who have been in a position to evaluate the leader's performance. While they do possess on-the-job observations of performance, they were not trained raters or assessors, thereby potentially affecting the reliability and subsequent validity of those ratings. Additionally, with regard to the Pulse survey, which serves as the follow up to the initial 360˚, often the leader being assessed has since been promoted or is managing a different team (and thus has different raters). As such, he/she may be rated against different standards than those of the initial survey. Further, as noted in Appendix H, for the Pulse survey item: *To what extent have you observed this individual make positive changes in his/her behavior over the last year?,* an asterisk beneath the ratings indicates that "some participants may receive a low rating on this question because they were previously seen as performing well and not in need of large changes," thereby potentially affecting the utility of the outcome measure (use of which generated no significant findings).

Further, using such an item as a performance indicator could also call into question the psychometric feasibility of using a single item to measure a given construct (in this case, observations of behavioral change). However, literature in both the organizational justice (Jordan & Turner, 2008) and job satisfaction (Dolbier, Webster, McCalister, Mallon, & Steinhardt, 2005) domains have found psychometric support for single-item measures of a broader construct. Jordan and Turner (2008) assessed the feasibility of measuring organizational justice and its

three dimensions (distributive, interactional, procedural justice) with a single item. The authors found that a single item was a reliable measure of distributive (α = .83) and interactional (α = .89) justice, though the reliability for procedural justice (α = .54), did not meet Nunally's (1978) suggested cutoff of .70. The authors found that the single item was as valid as a broader measure, in one instance (single item measuring procedural justice) serving as an even better predictor of an outcome measure (job satisfaction) than the full-scale measure, accounting for 25% of the variance as opposed to 14% with the full-scale measure. The authors conclude that the single item measure was shown to be both a reliable and valid alternative to a full-scale measure of organizational justice. A similar study was conducted regarding job satisfaction (Dolbier et al., 2005) with similar results demonstrating reliability (α = .90) of a single item measure (compared to α = .92 for the full-scale measure), and demonstrating significant convergence (r = .82) with the full-scale measure, as well. So, while using a single item measure as a DV in the current study may have limited some breadth of ratings, previous research has supported single-item scales as reliable and valid measures of broader constructs.

Further, in the current research, study 2 outcome scores were aggregated within rater groups when more than one rater was present (direct report and peer rater levels), creating a continuous, normally distributed variable thereby meeting the assumption of normality to be tested in multiple regression. However, as each subject only had one leader rating, those ratings remained on an ordinal scale. While visual inspection of the histograms, boxplots, and scatter plots produced by the regression output suggested that the ordinal leader scores generally were distributed against a normal curve, because they were not truly on a continuum, the violation of the regression assumption(s) of normality (and normality of residuals) could be called into question. One implication of this may have been reduced variability in the outcome measures, and therefore generating a false negative (type 2) error in results (especially in the leader outcomes approaching significance in their relationship with the simulation BFF scores).

Finally, limitations regarding the utilized sample(s) exist. Data collected to evaluate H_1 were collected from an online sample (independent of the organization represented in study 2),

which may reduce generalizability of the convergence demonstrated between the FOS and the WLM-BFF to the sample in study 2. Additionally, the sample that was utilized for study 2 ($H_{2\text{-}4}$) was from a single organization, potentially reducing the generalizability of results to the broader leadership population across other organizations (with differing feedback environments). While using an organizationally homogenous sample may affect generalizability, doing so also controlled for the effects of feedback environment from an organization-wide perspective. While feedback environments may differ across functions or departments, each participant in the study 2 sample would have experienced an identical feedback environment at the organizational level. One area for future research could be to replicate the current study, though with measurements in place (and controlled for) for differences in feedback environment, perhaps through the use of Steelman, Levy, and Snell's (2004) Feedback Environment Scale.

CHAPTER 10

CONCLUSION AND APPLIED NEXT STEPS

The current research opens a number of doors for further investigation into the construct of behavioral feedback focus and feedback orientation. The evaluation of personality, cognitive, and assessment-related correlates of the construct could contribute to an even deeper understanding of "what it is and why it matters." While the current research proposed even more questions than it sought to, or did address, this research also lays a compelling foundation for a number of tactical and tangible next steps, which can be incorporated into LWF's, and the broader field's, consulting tools, processes, and methods immediately.

The results of this study sought to address two primary research questions: a) Is a construct similar to feedback orientation implicitly measured by the Worthy Leadership Model (i.e., WLM-BFF)? and b) Does the WLM-BFF predict leadership (work-related) outcomes as rated by others? Results indicated that with regard to the first research question, indeed a feedback orientation-like construct is implicitly measured by the Worthy Leadership Model. The WLM-BFF did demonstrate convergence with the FOS, however, future research should evaluate the WLM-BFF and FOS by way of a confirmatory factor analysis to evaluate if the single component extracted in study one is indeed distinct from the one (or four) factors of the FOS.

With regard to the second research question, the results of study 2 suggested that unlike previous investigations into feedback orientation, the WLM-BFF as rated by others (assessors) did predict leadership and performance outcomes (H_2), as well as post-assessment change (H_3), and post-hoc analyses suggested that differences in variability in outcomes can be accounted for through different methods of composite scale construction (in this case, within and between simulations). The method by which this study was conducted may have implications over and above its findings, as well, as the current methodology enabled the researcher to evaluate a construct unrelated to the Worthy Leadership Model, implicate it as a composite scale, identify

ratings of that construct within pre-existing work simulations, and use that information to predict relevant and meaningful outcomes.

Utilizing a similar process could enable both researchers and practitioners to incorporate the measurement of constructs of specific interest to various clients, without compromising the integrity or consistency of the assessment process. Prior to 2014, the vision statement for Leadership Worth Following, LLC was *Developing Worthy Leadership through Great Consulting and Great Science.* The commitment to science at LWF, and the broader field of consulting psychology (Division 13 of the APA), is what truly differentiates Consulting and I/O Psychology from management consulting and/or self-proclaimed "life coaches." For example, LWF's website states, "Like our clients, LWF believes that effective, sustainable leadership strategies need to be based on more than personal opinion, folk wisdom or good intentions. For that reason, we employ a science-based methodology of 'test, evaluate, and improve' at every step of our process" (Worthyleadership.com, 2015). Thus, in addition to fulfilling academic requirements of the researcher, the current research sought to concurrently test, evaluate, and ultimately improve assessment processes used within the fields of I/O and Consulting Psychology, as well. Based on that commitment to science, a "start, stop, continue" evaluation of the assessment components and methodologies implicated in both the current research and the broader field is as follows.

10.1 Start

With regard to the "start" element, four suggestions that practitioners may benefit from taking into consideration would be around methods of capturing data at the leader-rater level, evaluating a method or process to maximize consistency across raters from a TI to a T2 360˚ survey, evaluating protocols for missing data, and providing frame of reference training to 360˚ raters. While only one effect was initially found at the leader rater level (inbox BFF related to leader ratings of worthy leadership; H_2), this is not to say that leaders are not in a position to evaluate performance and leadership objectively. One reason this finding may have occurred was due to the aforementioned lower variability (and more leptokurtic distributions) in leader ratings. If the lack of a relationship between BFF and leader ratings (compared to effects at direct report

and peer rater levels) is due to measurement error, or bias, one question to consider may be how to more accurately assess a leader's perspective on his or her employee's performance or leadership. Just as Braddy et al. (2013) suggest that some aspects of feedback orientation may be better assessed by others (rather than self-ratings), perhaps feedback on performance at the leader-rater level would be better assessed qualitatively, and rated by an assessor quantitatively. This may help limit halo and other rater errors, where leaders may be less willing or inclined to provide negative ratings despite being willing and open to discuss negative performance aspects in a more qualitative method (interviews, pre-calls, etc.). Further, there may have been a range restriction issue at play with regard to the leader ratings. The sample from which study 2 was evaluated was selected to participate in the development assessment process based on prior high performance, as rated by their leader. As a result, the sample is already relatively homogeneous, and performance somewhat centered around a higher than average performance mean (explaining why the lowest score on the leader rated performance item was a 3).

The second "start" piece to most effectively evaluate change using 360˚ ratings as an outcome may be around attempting to maximize consistency of rater groups from T1 to T2 administrations of a 360˚ feedback tool. While two of the three hypotheses of study 2 were partially supported at both the direct report and peer level, perhaps those effects would be even more pronounced should additional "controls" be in place. One limitation of this study, and the use of T2 360˚ as a criterion, is that the individuals being assessed by those tools may have been promoted, changed jobs, or otherwise be in a position to where their rater groups are dissimilar to what they were during the initial or baseline 360˚. Had leaders been promoted, perhaps they were then being held to a higher performance standard, despite improving overall in performance (which may look like a decrease when rated). Or, different individuals may be composing the peer and direct report rater groups between the T1 and T2 administration. If a way to control for these extraneous variables could be incorporated, for example by utilizing ratings where consistency in rater groups was present (and interrater reliability within rater groups examined), this may help practitioners obtain more reliable data as they relate to both performance and post-

assessment/developmental intervention change, as well as provide for an area of meaningful future research in differences in perceptions of change in performance and leadership in relation to the consistency of rater groups from T1 to T2.

Third, practitioners may evaluate the process by which behavioral data is both elicited and rated in an assessment process. This is a thread that must be carefully pulled, however, as inconsistent results across work simulations are often both valid and useful data, and would not indicate a lack of reliability in the assessment process, but rather the simulations are working as designed to capture different elements of leadership performance (writing skills in the inbox versus oral communication skills in the leader meeting). Ongoing calibration training can only seek to improve the interrater reliability of behavioral measurement, and therefore emphasize and capture to an even greater extent meaningful differences across work simulations.

Finally, as the above may relate to the broader issue of using 360˚ survey data as a criterion, administrators of 360˚ surveys may benefit from providing a brief frame of reference training session to each rater who will be participating in a given leader's 360˚ survey. This may help limit rater error (such as an unwillingness to rate someone objectively or negatively) as well as increase reliability across raters and rater groups.

10.2 Stop

One area for further consideration within executive assessment may be around the "super IV" principle of ratings within and between work simulations. Should a "super IV" methodology be explored further, it may warrant further investigation into behavioral rating "pruning" within each work simulation should a more targeted set of behaviors be found to be better or more reliably demonstrated and assessed in one or some work simulations compared to others. Doing so may allow for a more efficient honing of behavioral ratings within the simulation, perhaps at the expense, however, of capturing the breadth of the relevant competency model (an issue of quantity vs. quality of data generated by the process). Future research could investigate the pros and cons of rating behavioral work simulations against broad models or pruned and honed versions in the spirit of "super IVs." So, while outside of the true scope of "stop" with

regard to assessment methodology, these evaluations may challenge practitioners to incorporate a similar multi-rater multi-method methodology (which most do) to further explore the improvement of the reliability of their process and ultimately the utility and scientific and legal defensibility of both their predictors and outcomes.

10.3 Continue

With regard to LWF's assessment process specifically, a number of strengths that may be leveraged both at LWF and within the field more broadly were also identified during the course of the current research. First, the breadth of the WLM may allow for additional studies in similar process to be conducted, to meet client requests and needs on an even broader scale (regardless of their initial interest in "what" to assess). LWF's approach of assessing leadership through "four sets of eyes" allows for future research to be conducted leveraging a number of different perspectives and quantitative approaches. The ability to pull and aggregate data from testing (personality and cognitive), assessor ratings (from work simulations), self-ratings (from the 360˚ and structured interview), and organizational ratings (from the 360˚ rater groups), enables future research utilizing LWF's archival data and may help benefit and contribute to the broader field(s) of leadership development, executive assessment and coaching, and consulting psychology. And, as uncovered in the construction and subsequent testing of a super IV across work simulations, LWF, and other practitioners within the field of Consulting Psychology, will likely benefit from continuing to evaluate results (data) within individual measures ("down"), as well as across them, truly aggregating data both "across" and "down" to account for the most variability in the outcome of interest, as possible.

Further, while there is still some question as to the utility of measuring feedback orientation in a self-report format (Braddy et al., 2013), multi-rater, multi-method processes (whether at LWF or elsewhere) may also allow for an empirical evaluation of this question specifically as it relates to behavioral feedback focus. As behaviors are rated in a self-rating format, as well as "other" ratings captured in 360˚ administration, this may allow for a deeper

investigation into the implications of discrepancies between self and other ratings, especially as they relate to the feedback orientation (and WLM) dimension of self-awareness.

10.4 Application and Implications to the Broader Field

The current research may suggest an opportunity for continued investigation into the WLM-BFF as a standalone measurement tool to assess feedback seeking and receptivity behaviors in executive assessment. Should these ten behaviors stand up to further, confirmatory evaluation as to their reliability and validity in predicting performance (evaluated in H_2) as well as coachability and likelihood to develop (evaluated in H_3), they may be leveraged across industries, firms, and roles as a standalone construct to help organizations identify and develop the next generation(s) of employees, whether leaders or individual contributors. Having empirically evaluated behavioral feedback focus within the context of executive coaching also addresses an outstanding need for more rigorous evaluation of executive coaches, as well as the tools and processes they use. While individuals (such as Dale Thomson, 1987) and organizations (such as the Harvard Institute of Coaching) have attempted to bring rigor to the field of executive coaching and its evaluation of practitioners and outcomes, an opportunity still exists to apply the scientific rigor mandated by the field of psychology to the practice that "mobilizes strengths and realizes the potential of individuals and organizations" (Harvard IOC, 2015) as well as to the broader field of consulting psychology, which seeks to "apply psychological insights for the success and fulfillment of individuals, groups, and organizations" (Society of Consulting Psychology, 2015). It is the hope of this researcher that the current study accomplished just that – truly bridging the gap between rigorous empirically-based scientific investigation and the applied practice of executive assessment and coaching, as well as the broader field of consulting psychology.

Appendix A

Tables

Table 1. Demographics of Study 1 Survey Sample.

Variable	Frequency	Percent	*M*	*SD*
Age			27.71	14.70
Ethnicity				
American Indian or Alaskan Native	3	2.30		
Asian or Pacific Islander	5	3.90		
Black or African American	10	7.80		
Hispanic or Latino	10	7.80		
White/Caucasian	107	83.60		
Prefer not to Answer	1	.80		
Other	1	.80		
Highest Level of Education				
Some high school	2	1.56		
High school diploma	10	7.81		
Trade technical training	5	3.91		
Some college - no degree	15	11.72		
Associate degree	14	10.94		
Bachelor's degree	55	42.97		
Master's degree	22	17.19		
Professional degree - MD	4	3.13		
Doctorate - PhD	1	0.78		
Principal Industry				
Automotive	3	2.34		
Business Support & Logistics	17	13.28		
Construction, Machinery, and Homes	6	4.69		
Education	13	10.16		
Entertainment & Leisure	5	3.91		
Finance & Financial Services	9	7.03		
Food & Beverages	6	4.69		
Government	4	3.13		
Healthcare & Pharmaceuticals	9	7.03		
Insurance	2	1.56		
Manufacturing	12	9.38		
Nonprofit	8	6.25		
Retail & Consumer Durables	9	7.03		
Real Estate	4	3.13		
Telecommunications, Technology, Internet & Electronics	7	5.47		
Transportation & Delivery	2	1.56		
I am currently not employed	12	9.38		
Level of Leadership				
Manager/Supervisor	46	35.94		
Sr. Manager/Director	21	16.41		
Executive (VP, SVP)	13	10.16		
C-Level (CEO, COO, CFO)	48	37.50		
Number of Direct Reports			9.71	8.93

Table 2. Worthy Leadership Model - Behavioral Feedback Focus (WLM-BFF) - Item & Item-Total Statistics.

	WLM-BFF Item	*M*	*SD*	Corrected Item-Total Correlation	Chronbach's Alpha if Item Deleted
1.	Encourages others to voice their needs and concerns in potentially adversarial situations.	3.61	1.09	.61	.89
2.	Can hear bad news and negative feedback without becoming defensive.	3.77	.87	.63	.89
3.	Listens carefully and accurately discerns others' needs and concerns.	4.05	.85	.75	.88
4.	Establishes an environment in which the ideas and suggestions of others are encouraged.	4.16	.84	.73	.88
5.	Demonstrates a keen awareness of own motives, behavior, and impact on others.	4.07	.80	.69	.88
6.	Seeks opportunities to learn about self and elicit feedback from others.	3.82	.85	.71	.88
7.	Translates developmental feedback and lessons of experience into action.	3.88	.88	.67	.88
8.	Seeks new challenges and opportunities for personal and professional growth.	4.04	.89	.64	.89
9.	Acknowledges and takes responsibility for own mistakes.	4.32	.77	.55	.89
10.	Accepts praise graciously.	3.83	.87	.48	.90

Table 3. Feedback Orientation Scale Items & Item-Total Statics.

Dimension/Item	***M***	***SD***	**Corrected Item-Total Correlation**	**Chronbach's Alpha if Item Deleted**
Utility				
1. Feedback contributes to my success at work.	4.13	.89	.69	.93
2. To develop my skills at work, I rely on feedback.	3.94	.87	.68	.93
3. Feedback is critical for improving performance.	4.20	.76	.71	.93
4. Feedback from supervisors can help me advance in a company.	4.04	.92	.64	.93
5. I find that feedback is critical for reaching my goals.	4.13	.84	.65	.93
Accountability				
1. It is my responsibility to apply feedback to improve my performance.	4.22	.77	.64	.93
2. I hold myself accountable to respond to feedback appropriately.	4.23	.67	.55	.93
3. I don't feel a sense of closure until I respond to feedback.	3.56	.96	.47	.93
4. If my supervisor gives me feedback, it is my responsibility to respond to it.	3.96	.85	.62	.93
5. I feel obligated to make changes based on feedback.	3.91	.87	.53	.93
Social Awareness				
1. I try to be aware of what other people think of me.	3.98	.82	.50	.93
2. Using feedback, I am more aware of what people think of me.	3.98	.80	.67	.93
3. Feedback helps me manage the impression I make on others.	4.05	.76	.74	.93
4. Feedback lets me know how I am perceived by others.	4.08	.86	.67	.93
5. I rely on feedback to help me make a good impression.	3.82	.92	.65	.93
Feedback Self-Efficacy				
1. I feel self-assured when dealing with feedback.	3.76	.93	.69	.93
2. Compared to others, I am more competent at handling feedback.	3.73	.85	.60	.93
3. I believe that I have the ability to deal with feedback effectively.	4.11	.71	.66	.93
4. I feel confident when responding to both positive and negative feedback.	4.02	.78	.51	.93
5. I know that I can handle the feedback that I receive.	4.13	.79	.58	.93

Table 4. WLM-BFF & FOS Factor (Dimension) - Level Correlations and Reliability.

Variable (Dimension/Full Scale)	*α*	WLM-BFF Capacity	WLM - BFF Commitment	WLM - BFF Character	WLM-BFF	FOS - Utility	FOS - Accountability	FOS - Self-Awareness	FOS - Self-Effacy
WLM-BFF Capacity	0.63								
WLM - BFF Commitment	0.88	.71**							
WLM - BFF Character	0.55	.49**	.58**						
WLM-BFF	0.89	.83**	.96**	.73**					
FOS - Utility	0.87	.35**	.56**	.47**	.55**				
FOS - Accountability	0.73	.31**	.47**	.40**	.47**	.70**			
FOS - Self-Awareness	0.85	.35**	.49**	.47**	.51**	.70**	.73**		
FOS - Self-Effacy	0.87	.50**	.56**	.54**	.61**	.59**	.61**	.57**	
FOS	0.93	.44**	.61**	.55**	.63**	.88**	.87**	.87**	.81**

** $p < .001$

Table 5. Demographics of Study 2 Sample.

Variable	Frequency	Percent	*M*	*SD*
Age			38.09	6.13
Gender				
Male	94	79.70		
Female	24	20.30		
Ethnicity				
African	9	7.63		
Asian	1	0.85		
Caucasian	95	80.51		
Latino	12	10.17		
Other	1	0.85		
Highest Level of Education				
Some college - no degree	10	8.47		
Associate degree	3	2.54		
Bachelor's degree	57	48.31		
Master's degree	38	32.20		
Professional degree - MD	8	6.78		
Doctorate - PhD	2	1.69		

Table 6. Descriptive Statistics of Independent Variables.

Variables (IV)	***α***	***N***	**Min**	**Max**	***M***	***SD***	**Skewness**	***SE***	**Kurtosis**	***SE***
Inbox BFF	.68	118	2.25	3.81	3.10	.32	-.03	.22	-.51	.44
Team Meeting BFF	.92	118	1.67	4.50	2.96	.54	.48	.22	.79	.44
Leader Meeting BFF	.80	118	2.00	4.70	3.00	.39	.52	.22	2.83	.44

Table 7. Correlations between Independent Variables.

Variables	Inbox BFF	Team Meeting BFF
Inbox BFF	1	
Team Meeting BFF	0.01	1
Leader Meeting BFF	0.12	0.15

Table 8. Descriptive Statistics of Control Variables from T1 360 Survey (Baseline).

Variables (Control)	***N***	**Min**	**Max**	***M***	***SD***	**Skewness**	***SE***	**Kurtosis**	***SE***
Performance - Direct Reports	97	2.71	5.00	4.16	.53	-.73	.24	.40	.49
Performance - Peers	114	2.50	5.00	4.01	.44	-.74	.23	1.79	.45
Performance - Leader	115	3.00	5.00	4.12	.53	.12	.23	.39	.45
Leadership - Direct Reports	97	3.00	5.00	4.32	.54	-.85	.24	.22	.49
Leadership - Peers	114	2.50	5.00	4.10	.45	-.29	.23	.87	.45
Leadership - Leader	115	3.00	5.00	4.23	.60	-.13	.23	-.45	.45

Table 9. Correlations between Control (Baseline) Variables.

Variables	Baseline Leadership Direct Reports	Baseline Performance Direct Reports	Baseline Leadership Peers	Baseline Performance Peers	Baseline Leadership Leader	Baseline Performance Leader
Baseline Leadership - Direct Reports	1.00					
Baseline Performance - Direct Reports	.76**	1.00				
Baseline Leadership - Peers	.17	.27**	1.00			
Baseline Performance - Peers	.08	.19	.58**	1.00		
Baseline Leadership - Leader	.04	.09	.15	.15	1.00	
Baseline Performance - Leader	-.05	.12	.20*	.21*	.35**	1.00

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

Table 10. Descriptive Statistics of H_2 Dependent Variables.

Variables (DVs)	*N*	Min	Max	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Performance - Direct Reports	104	2.67	5.00	4.18	.53	-.09	.24	-.47	.47
Performance - Peers	111	2.33	5.00	3.96	.48	-.45	.23	1.21	.46
Performance - Leader	117	3.00	5.00	3.97	.55	-.02	.22	.41	.44
Leadership - Direct Reports	104	2.25	5.00	4.15	.54	-.38	.24	.53	.47
Leadership - Peers	112	2.00	5.00	3.96	.47	-.67	.23	2.43	.45
Leadership - Leader	118	2.00	5.00	3.96	.53	-.40	.22	1.90	.44

Table 11. Descriptive Statistics of H_3 Dependent Variables - Change (Difference) Scores from Baseline (T1) to Pulse Survey (T2).

Variables (DVs)	*N*	Min	Max	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Change in Performance - Direct Reports	89	-1.33	1.50	.02	.58	.44	.26	.28	.51
Change in Performance - Peers	107	-1.50	1.50	-.05	.56	.24	.23	.55	.46
Change in Performance - Leader	114	-2.00	1.00	-.15	.69	-.11	.23	-.22	.45
Change in Leadership - Direct Reports	89	-1.33	1.33	-.13	.56	.23	.26	-.18	.51
Change in Leadership - Peers	108	-1.75	1.50	-.13	.57	.36	.23	1.13	.46
Change in Leadership - Leader	115	-3.00	1.00	-.29	.72	-.36	.23	.91	.45

Table 12. Descriptive Statistics of H4 Dependent Variables - Observed Behavioral Change.

Variables (DVs)	*N*	Min	Max	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Observed Behavioral Change - Direct Reports	104	2.00	5.00	3.38	.63	.14	.24	-.36	.47
Observed Behavioral Change - Peers	111	1.50	4.25	3.24	.55	-.61	.23	.31	.46
Observed Behavioral Change - Leader	113	1.00	5.00	3.68	.77	-.21	.23	.40	.45

Table 13. Correlations between dependent variables.

Variables	Performance - Direct Reports	Performance - Peers	Performance - Leader	Leadership - Direct Reports	Leadership - Peers	Leadership - Leader	Change in Performance - Direct Reports	Change in Performance - Peers	Change in Performance - Leader	Change in Leadership - Direct Reports	Change in Leadership - Peers	Change in Leadership - Leader	Observed Behavioral Change - Direct Reports	Observed Behavioral Change - Peers	Observed Behavioral Change - Leader
Performance - Direct Reports	1.00														
Performance - Peers	.43**	1.00													
Performance - Leader	.17	.20*	1.00												
Leadership - Direct Reports	.82**	.40**	.16	1.00											
Leadership - Peers	.32**	.74**	.33**	.35**	1.00										
Leadership - Leader	.18	.14	.56**	.16	.20*	1.00									
Change in Performance - Direct Reports	.54**	.21	.05	.36**	.23*	.03	1.00								
Change in Performance - Peers	.24*	.64**	.00	.24*	.42**	-.12	.23*	1.00							
Change in Performance - Leader	.13	.10	.66**	.11	.21*	.32**	.09	.08	1.00						
Change in Leadership - Direct Reports	.40**	.23*	.13	.46**	.22*	.15	.71**	.25*	.00	1.00					
Change in Leadership - Peers	.11	.39**	.13	.06	.65**	-.06	.23*	.54**	.19	.12	1.00				
Change in Leadership - Leader	.09	.03	.33**	.11	.18	.58**	.07	-.02	.39**	.15	.08	1.00			
Observed Behavioral Change - Direct Reports	.42**	.03	.20*	.39**	.14	.17	.40**	.04	.19	.39**	.12	.07	1.00		
Observed Behavioral Change - Peers	.17	.34**	.10	.16	.42**	.06	.10	.18	.08	.05	.31**	-.02	.04	1.00	
Observed Behavioral Change - Leader	.10	.07	.30**	-.03	.10	.24**	.09	-.02	.16	.15	.01	.23*	.06	.06	1.00

* $p < .05$

* *$p < .01$

Table 14. Standard Multiple Regression of Simulation BFF scores on Pulse Survey Performance at the Direct Report Rater Level.

Variables	Performance Direct Reports (DV)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Inbox BFF	.19				.36*	.16	.04	.67	.22	.05
Team Meeting BFF	.08	-.03			.10	.09	-.08	.29	.11	.01
Leader Meeting BFF	-.11	.14	.16		-.21	.14	-.48	.06	-.16	.02
			Intercept = 3.42							
Means	4.18	3.10	2.96	3.01					R^2 = .07	
									Adjusted R^2 = .04	
Std. Deviations	.53	.32	.56	.39					*R* = .26	

* $p < .05$

Table 15. Standard Multiple Regression of Simulation BFF scores on Pulse Survey Performance at the Peer Rater Level.

Variables	Pulse Performance Peers (DV)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Inbox BFF	.19				.30*	.14	.03	.58	.21	.04
Team Meeting BFF	.08	.05			.08	.09	-.09	.25	.09	.01
Leader Meeting BFF	-.12	.11	.15		-.20	.12	-.43	.04	-.16	.02
				Intercept = 3.36						
Means	3.96	3.11	2.94	3.01						R^2 = .07
										Adjusted R^2 = .04
Std. Deviations	.48	.33	.53	.39						*R* = .26

* $p < .05$

Table 16. Standard Multiple Regression of Simulation BFF scores on Pulse Survey Performance at the Leader Rater Level.

Variables	Pulse Performance Leader (DV)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Inbox BFF	.10				.14	.16	-.18	.45	.08	.01
Team Meeting BFF	.06	.00			.04	.09	-.14	.23	.04	.00
Leader Meeting BFF	.14	.12	.15		.18	.13	-.09	.44	.13	.02
				Intercept = 2.89						
Means	3.97	3.10	2.96	3.00						R^2 = .03
										Adjusted R^2 = .002
Std. Deviations	.55	.32	.55	.39						R = .17

Table 17. Standard Multiple Regression of Simulation BFF scores on Pulse Survey Leadership at the Direct Report Rater Level.

Variables	Leadership Direct Reports (DV)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Inbox BFF	.09				.20	.16	-.12	.52	.12	.01
Team Meeting BFF	.16	.05			.19*	.09	.01	.38	.20	.04
Leader Meeting BFF	-.15	.10	.17		-.27	.14	-.54	.00	-.19	.04
				Intercept = 3.78						
Means	4.15	3.09	2.96	3.00						R^2 = .07
										Adjusted R^2 = .04
Std. Deviations	.54	.32	.56	.39						R = .27

* $p < .05$

Table 18. Standard Multiple Regression of Simulation BFF scores on Pulse Survey Leadership at the Peer Rater Level.

Variables	Leadership Peers (DV)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Inbox BFF	.23				.34*	.13	.08	.61	.24	.06
Team Meeting BFF	.06	.05			.06	.08	-.11	.22	.07	.00
Leader Meeting BFF	-.06	.10	.17		-.11	.11	-.34	.11	-.10	.01
				Intercept = 3.07						
Means	3.96	3.11	2.93	3.00						R^2 = .07
										Adjusted R^2 = .04
Std. Deviations	.47	.33	.54	.39						*R* = .26

* $p < .05$

Table 19. Standard Multiple Regression of Simulation BFF scores on Pulse Survey Leadership at the Leader Rater Level.

Variables	Leadership Leader (DV)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Inbox BFF	.20				.30*	.15	.00	.60	.18	.03
Team Meeting BFF	.07	.01			.06	.09	-.12	.23	.06	.00
Leader Meeting BFF	.12	.12	.15		.13	.13	-.13	.38	.09	.01
				Intercept = 2.47						
Means	3.96	3.10	2.96	3.00						R^2 = .05
										Adjusted R^2 = .03
Std. Deviations	.53	.32	.54	.39						R = .23

* $p < .05$

Table 20. Standard Multiple Regression of Simulation BFF scores and Baseline Performance on Change in Performance at the Direct Report Rater Level.

Variables	Change in Performance Direct Reports (DV)	Baseline Performance Direct Reports (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance Direct Reports (Control)	-.59					-.67**	.10	-.86	-.49	-.61	.35
Inbox BFF	.16	.08				.41*	.16	.10	.72	.23	.05
Team Meeting BFF	-.04	.10	-.05			.05	.09	-.13	.23	.05	0
Leader Meeting BFF	-.11	.08	.12	.13		-.17	.15	-.45	.12	-.10	0
					Intercept = 1.89						
Means	.02	4.16	3.08	2.99	2.98						R^2 = .40
											Adjusted R^2 = .37
Std. Deviations	.58	.52	.32	.55	.34						R = .63

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

Table 21. Standard Multiple Regression of Simulation BFF scores and Baseline Performance on Change in Performance at the Peer Rater Level.

Variables	Change in Performance Peers (DV)	Baseline Performance Peers (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance - Peers (Control)	-.55					-.68**	.10	-.88	-.48	-.55	.29
Inbox BFF	.15	.03				.31*	.14	.04	.59	.18	.03
Team Meeting BFF	.05	.05	.06			.09	.09	-.08	.26	.09	.01
Leader Meeting BFF	-.18	.12	.13	.09		-.21	.12	-.44	.03	-.14	.02
					Intercept = 2.01						
Means	-.05	4.01	3.11	2.93	3.01						R^2 = .36
											Adjusted R^2 = .33
Std. Deviations	.56	.44	.32	.52	.39						R = .60

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

Table 22. Standard Multiple Regression of Simulation BFF scores and Baseline Performance on Change in Performance at the Leader Rater Level.

Variables	Change in Performance Leader (DV)	Baseline Performance Leader (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance - Leader (Control)	-.62					-.84**	.10	-1.04	-.63	-.64**	.38
Inbox BFF	-.08	.20				.09	.17	-.24	.42	.04	.00
Team Meeting BFF	.07	-.20	.01			.06	.10	-.13	.25	.05	.00
Leader Meeting BFF	-.05	.22	.14	.09		.14	.14	-.13	.42	.08	.01
					Intercept = 2.42						
Means	-.15	4.12	3.09	2.95	3.00						R^2 = .39
											Adjusted R^2 = .37
Std. Deviations	.69	.53	.32	.53	.38						R = .63

** $p < .001$

Table 23. Standard Multiple Regression of Simulation BFF scores and Baseline Leadership on Change in Leadership at the Direct Report Rater Level.

Variables	Change in Leadership Direct Reports (DV)	Baseline Leadership Direct Reports (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership Direct Reports (Control)	-.61					-.63**	.09	-.80	-.46	-.62	.37
Inbox BFF	.18	-.05				.29	.12	-.01	.58	.16	.03
Team Meeting BFF	.03	.14	-.05			.13	.09	-.04	.30	.13	.02
Leader Meeting BFF	-.09	.09	.12	.13		-.11	.14	-.38	.17	-.07	.00
					Intercept = 1.63						
Means	-.13	4.10	3.08	2.99	2.98						R^2 = .42
											Adjusted R^2 = .39
Std. Deviations	.56	.45	.32	.55	.34						*R* = .65

** $p < .001$

Table 24. Standard Multiple Regression of Simulation BFF scores and Baseline Leadership on Change in Leadership at the Peer Rater Level.

Variables	Change in Leadership Peers (DV)	Baseline Leadership Peers (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership Peers (Control)	-.60					-.76**	.10	-.96	-.56	-.59	.34
Inbox BFF	.20	-.01				.36*	.14	.09	.63	.20	.04
Team Meeting BFF	.01	.09	.06			.07	.09	-.10	.24	.06	.00
Leader Meeting BFF	-.12	.13	.12	.11		-.12	.11	-.34	.11	-.08	.01
					Intercept = 2.02						
Means	-.13	4.10	3.11	2.92	3.00						R^2 = .40
											Adjusted R^2 = .38
Std. Deviations	.57	.45	.32	.52	.39						R = .64

* $p < .05$

** $p < .001$

Table 25. Standard Multiple Regression of Simulation BFF scores and Baseline Leadership on Change in Leadership at the Leader Rater Level.

Variables	Leadership Leader (DV)	Baseline Leadership Leader (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership Leader (Control)	-.70					-.87**	.08	-1.04	-.71	-.72	.49
Inbox BFF	.05	.09				.24	.15	-.07	.54	.11	.01
Team Meeting BFF	-.10	.21	.01			.05	.09	-.14	.24	.04	.00
Leader Meeting BFF	.03	.09	.14	.09		.16	.13	-.10	.41	.08	.01
					Intercept = 2.06						
Means	-.29	4.24	3.09	2.95	3.00						R^2 = .51
											Adjusted R^2 = .49
Std. Deviations	.72	.60	.32	.53	.38						R = .71

** $p < .001$

Table 26. Standard Multiple Regression of Simulation BFF scores and Baseline Performance on Pulse Survey Observed Behavioral Change at the Direct Report Rater Level.

Variables	Observed Behavioral Change Direct Reports (DV)	Baseline Performance Direct Reports (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance Direct Reports (Control)	-.07					-.09	.13	-.34	.17	-.07	.01
Inbox BFF	.08	.08				.17	.21	-.25	.59	.09	.01
Team Meeting BFF	-.07	.11	-.05			-.07	.12	-.31	.18	-.06	.00
Leader Meeting BFF	-.02	.08	.12	.13		-.04	.20	-.43	.35	-.02	.00
					Intercept = 3.55						
Means	3.40	4.16	3.08	2.99	2.98						R^2 = .02
											Adjusted R^2 = -.03
Std. Deviations	.62	.52	.32	.55	.34						*R* = .13

Table 27. Standard Multiple Regression of Simulation BFF scores and Baseline Performance on Pulse Survey Observed Behavioral Change at the Peer Rater Level.

Variables	Observed Behavioral Change Peers (DV)	Baseline Performance Peers (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance - Peers (Control)	.13					.14	.12	-.09	.38	.12	.01
Inbox BFF	.19	.03				.31	.17	-.02	.63	.18	.03
Team Meeting BFF	.09	.05	.05			.08	.10	-.13	.28	.07	.00
Leader Meeting BFF	.10	.11	.13	.12		.08	.14	-.20	.35	.05	.00
					Intercept = 1.25						
Means	3.23	4.01	3.11	2.92	3.00						R^2 = .06
											Adjusted R^2 = .03
Std. Deviations	.55	.45	.32	.52	.39						*R* = .25

Table 28. Standard Multiple Regression of Simulation BFF scores and Baseline Performance on Pulse Survey Observed Behavioral Change at the Leader Rater Level.

Variables	Observed Behavioral Change Leader (DV)	Baseline Performance Leader (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance - Leader (Control)	.10					.13	.15	-.16	.41	.09	.01
Inbox BFF	.03	.22				.02	.25	-.47	.51	.01	.00
Team Meeting BFF	-.01	-.02	-.01			-.02	.14	-.30	.27	-.01	.00
Leader Meeting BFF	.07	.22	.14	.07		.10	.20	-.31	.49	.05	.00
					Intercept = 2.86						
Means	3.67	4.13	3.08	2.96	3.00						R^2 = .01
											Adjusted R^2 = -.03
Std. Deviations	.78	.54	.32	.53	.38						R = .11

Table 29. Standard Multiple Regression of Simulation BFF scores and Baseline Leadership on Pulse Survey Observed Behavioral Change at the Direct Report Rater Level.

Variables	Observed Behavioral Change Direct Reports (DV)	Baseline Leadership Direct Reports (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership Direct Reports (Control)	-.05					-.04	.12	-.28	.21	-.03	.00
Inbox BFF	.08	-.05				.16	.21	-.27	.58	.08	.01
Team Meeting BFF	-.07	.14	-.05			-.07	.12	-.31	.18	-.06	.00
Leader Meeting BFF	-.02	.09	.12	.13		-.04	.20	-.44	.35	-.02	.00
					Intercept = 3.41						
Means	3.41	4.32	3.08	2.99	2.98						R^2 = .01
											Adjusted R^2 = -.03
Std. Deviations	.62	.55	.32	.55	.34						R = .12

Table 30. Standard Multiple Regression of Simulation BFF scores and Baseline Leadership on Pulse Survey Observed Behavioral Change at the Peer Rater Level.

Variables	Observed Behavioral Change Peers (DV)	Baseline Leadership Peers (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership Peers (Control)	.05					.05	.12	-.19	.29	.04	.00
Inbox BFF	.19	-.01				.31	.17	-.02	.64	.18	.03
Team Meeting BFF	.09	.09	.05			.08	.10	-.13	.28	.07	.00
Leader Meeting BFF	.10	.13	.13	.12		.08	.14	-.19	.36	.06	.00
					Intercept = 1.58						
Means	3.23	4.01	3.11	2.92	2.99						R^2 = .05
											Adjusted R^2 = .01
Std. Deviations	.55	.45	.32	.52	.39						*R* = .22

Table 31. Standard Multiple Regression of Simulation BFF scores and Baseline Leadership on Pulse Survey Observed Behavioral Change at the Leader Rater Level.

Variables	Observed Behavioral Change Leader (DV)	Baseline Leadership Leader (Control)	Inbox BFF	Team Meeting BFF	Leader Meeting BFF	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership Leader (Control)	-.07					-.10	.13	-.35	.16	-.08	.01
Inbox BFF	.03	.07				.07	.24	-.41	.55	.03	.00
Team Meeting BFF	-.01	.20	-.01			.00	.15	-.29	.29	.00	.00
Leader Meeting BFF	.07	.08	.07	.07		.14	.20	-.26	.53	.07	.00
					Intercept = 3.44						
Means	3.67	4.24	3.08	2.96	3.00						R^2 = .01
											Adjusted R^2 = -.03
Std. Deviations	.78	.60	.32	.53	.38						*R* = .10

Table 32. Frequencies of BFF Behavioral Level Outcome Relationships Across Work Simulations

BFF Behavior	1	2	3	4	5	6	7	8	9	10
Competency	Capacity	Capacity	Commit	Commit	Commit	Commit	Commit	Commit	Character	Character
IB	2	2	1	2	0	3	1	2	0	0
TM	0	1	0	0	0	2	0	1	1	1
LM	0	0	2	1	4	0	2	0	0	0
Selected Simulation	IB	IB	LM	IB	LM	IB	LM	IB	TM	TM

Table 33. Descriptive Statistics of Super BFF IV

Variables (DVs)	N	Min	Max	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
Super BFF IV	112	2.38	3.80	3.05	.28	.18	.23	-.34	.45

Table 34. Regression of Super IV BFF score on Performance at the Direct Report Rater Level.

Variables	Performance Direct Reports	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*
Super IV	.07		.13	.19	-.24	.51	.07
Mean	4.20	3.05			Intercept = 3.79		
Std. Deviation	.53	.28					

Table 35. Regression of Super IV BFF score on Performance at the Peer Rater Level.

Variables	Performance Peers	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*
Super IV	.06		.10	.16	-.22	.42	.06
Mean	3.98	3.06		Intercept = 3.69			
Std. Deviation	.46	.28					

Table 36. Regression of Super IV BFF score on Performance at the Leader Rater Level

Variables	Performance Leaders	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*
Super IV	.27		0.53**	.18	.17	.89	.27
Mean	3.99	3.05					
					Intercept = 2.37		
Std. Deviation	.55	.28					

** $p < .001$

Table 37. Regression of Super IV BFF score on Worthy Leadership at the Direct Report Rater Level

Variables	Worthy Leadership Direct Reports	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*
Super IV	.03		.06	.19	-.33	.44	.03
Mean	4.17	3.05					
					Intercept = 4.00		
Std. Deviation	.53	.28					

Table 38. Regression of Super IV BFF score on Worthy Leadership at the Peer Rater Level

Variables	Worthy Leadership Peers	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*
Super IV	.18		.27	.15	-.02	.57	.18
Mean	3.99	3.05					
					Intercept = 3.15		
Std. Deviation	.43	.28					

Table 39. Regression of Super IV BFF score on Worthy Leadership at the Leader Rater Level

Variables	Worthy Leadership Leader	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*
Super IV	.30		.56**	.17	.23	.89	.30
Mean	3.99	3.05			Intercept = 2.29		
Std. Deviation	.51	.28					

** $p < .001$

Table 40. Standard Multiple Regression of Super IV BFF scores and Baseline Performance on Change in Performance at the Direct Report Rater Level.

Variables	Change in Performance Direct Reports	Baseline Performance	Super IV	B	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance (Control)	-.58			-.65**	.10	-.85	-.45	-.60	.35
Super IV	.00	.16		.20	.19	-.18	.59	.10	.01
Means	.01	4.19	3.03						R^2 = .35
				Intercept = 2.12					Adjusted R^2 = .33
Std. Deviations	.56	.52	.27						R = .59

** $p < .001$

Table 41. Standard Multiple Regression of Super IV BFF scores and Baseline Performance on Change in Performance at the Peer Rater Level.

Variables	Change in Performance Peers	Baseline Performance	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance (Control)	-.56			-.70**	.11	-.91	-.50	-.57	.32
Super IV	-.02	.10		.08	.16	-.24	.40	.04	.00
Means	-.06	4.03	3.06						R^2 = .32
				Intercept = 2.54					Adjusted R^2 = .30
Std. Deviations	.54	.43	.28						R = .56

** $p < .001$

Table 42. Standard Multiple Regression of Super IV BFF scores and Baseline Performance on Change in Performance at the Leader Rater Level.

Variables	Change in Performance Leader	Baseline Performance	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance (Control)	-.63			-.90**	.10	-1.10	-.70	-.68	.44
Super IV	.02	.26		.50*	.19	.12	.88	.20	.04
Means	-.15	4.14	3.05						R^2 = .44
				Intercept = 2.05					Adjusted R^2 = .43
Std. Deviations	.71	.54	.28						R = .66

*p < .05

** p < .001

Table 43. Standard Multiple Regression of Super IV BFF scores and Baseline Leadership on Change in Worthy Leadership at the Direct Report Rater Level.

Variables	Change in Leadership Direct Report	Baseline Leadership	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership (Control)	-.64			-.66**	.09	-.84	-.49	-.65	.42
Super IV	.04	.09		.20	.18	-.16	.56	.09	.01
Means	-.12	4.34	3.03						R^2 = .42
				Intercept = 2.14					Adjusted R^2 = .40
Std. Deviations	.57	.55	.27						*R* = .65

** $p < .001$

Table 44. Standard Multiple Regression of Super IV BFF scores and Baseline Leadership on Change in Worthy Leadership at the Peer Rater Level.

Variables	Change in Leadership Peer	Baseline Leadership	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership (Control)	-.64			-.83**	.10	-1.02	-.63	-.65	.42
Super IV	.08	.09		.27	.15	-.03	.57	.14	.02
Means	-.13	4.12	3.06						R^2 = .43
				Intercept = 2.44					Adjusted R^2 = .41
Std. Deviations	.55	.43	.28						*R* = .65

** $p < .001$

Table 45. Standard Multiple Regression of Super IV BFF scores and Baseline Leadership on Change in Worthy Leadership at the Leader Rater Level.

Variables	Change in Leadership Leader	Baseline Leadership	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership (Control)	-.72			-.89**	.08	-1.05	-.73	-.74	.54
Super IV	.10	.13		.50*	.17	.16	.84	.19	.04
Means	-.28	4.26	3.05						R^2 = .55
				Intercept = 1.99					Adjusted R^2 = .54
Std. Deviations	.72	.60	.28						*R* = .74

* $p < .05$

** $p < .001$

Table 46. Standard Multiple Regression of Super IV BFF scores and Baseline Performance on Observed Behavioral Change at the Direct Report Rater Level.

Variables	Observed Behavioral Change Direct Reports	Baseline Performance	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance (Control)	-.09			-.10	.14	-.37	.17	-.08	.01
Super IV	-.04	.16		-.05	.26	-.58	.47	-.02	.00
Means	3.41	4.19	3.03	Intercept = 3.98					R^2 = .01 Adjusted R^2 = -.02
Std. Deviations	.63	.52	.27						R = .09

Table 47. Standard Multiple Regression of Super IV BFF scores and Baseline Performance on Observed Behavioral Change at the Peer Rater Level.

Variables	Observed Behavioral Change Peers	Baseline Performance	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance (Control)	.11			.12	.13	-.13	.38	.10	.01
Super IV	.19	.09		.38	.20	-.02	.77	.19	.04
Means	3.24	4.04	3.06						R^2 = .05
				Intercept = 1.59					Adjusted R^2 = .03
Std. Deviations	.56	.43	.28						*R* = .22

Table 48. Standard Multiple Regression of Super IV BFF scores and Baseline Performance on Observed Behavioral Change at the Leader Rater Level.

Variables	Observed Behavioral Change Leader	Baseline Performance	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Performance (Control)	.07			.06	.14	-.23	.34	.14	.00
Super IV	.11	.27		.27	.29	-.30	.84	.10	.01
Means	3.67	4.14	3.04						R^2 = .01
				Intercept = 2.62					Adjusted R^2 = -.01
Std. Deviations	.77	.55	.27						*R* = .11

Table 49. Standard Multiple Regression of Super IV BFF scores and Baseline Leadership on Observed Behavioral Change at the Direct Report Rater Level.

Variables	Observed Behavioral Change Direct Report	Baseline Leadership	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership (Control)	-.06			-.07	.13	-.32	.19	-.06	.00
Super IV	-.04	.09		-.07	.26	-.59	.45	-.03	.00
Means	3.41	4.34	3.03						R^2 = .004
				Intercept = 3.90					Adjusted R^2 = -.02
Std. Deviations	.63	.55	.27						*R* = .07

Table 50. Standard Multiple Regression of Super IV BFF scores and Baseline Leadership on Observed Behavioral Change at the Peer Rater Level.

Variables	Observed Behavioral Change Peer	Baseline Leadership	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership (Control)	.03			.02	.13	-.24	.27	.02	.00
Super IV	.20	.09		.39	.20	.00	.78	.20	.04
Means	3.24	4.12	3.05						R^2 = .04
				Intercept = 1.97					Adjusted R^2 = .02
Std. Deviations	.56	.43	.28						*R* = .20

Table 51. Standard Multiple Regression of Super IV BFF scores and Baseline Leadership on Observed Behavioral Change at the Leader Rater Level.

Variables	Observed Behavioral Change Leader	Baseline Leadership	Super IV	*B*	*SE*	95% Confidence Interval for *B* Lower	Upper	*β*	sr^2 (unique)
Baseline Leadership (Control)	-.07			-.10	.13	-.35	.15	-.08	.01
Super IV	.11	.11		.32	.28	-.23	.87	.12	.01
Means	3.67	4.26	3.04						R^2 = .02
				Intercept = 3.11					Adjusted R^2 = -.002
Std. Deviations	.77	.61	.27						*R* = .13

Appendix B

The Worthy Leadership Model (WLM)

The Worthy Leadership Model

The capacity to lead

1. Capacity to reason and make good decisions
   Critical thinking & decision making; numerical & financial acumen
2. Capacity to see & realize the future
   Seeing the future; strategic & execution excellence; global acumen
3. Capacity to communicate & influence
   Inspiration & influence; conflict management; leadership presence
4. Capacity to know
   Job & industry knowledge; business knowledge
5. Capacity to persevere & adapt
   Energy, adaptability, & humor

The commitment to lead

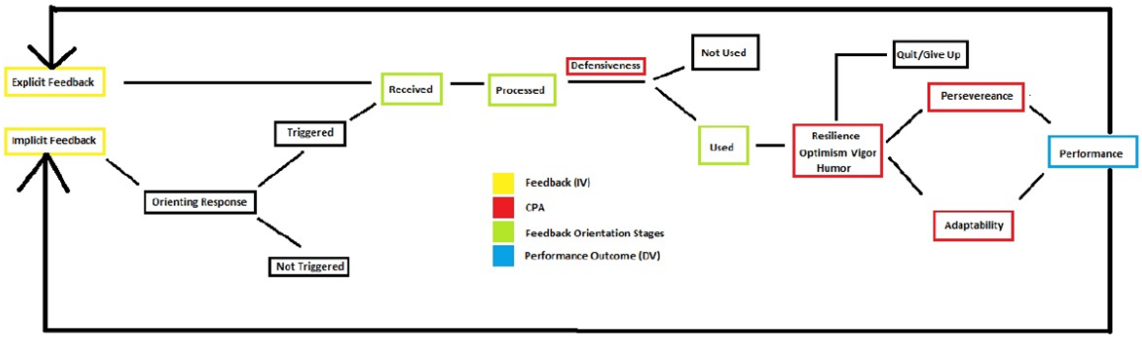
6. Commitment to excellence
   Defines success; passion for results
7. Commitment to people & relationships
   Building talent; diversity & culture; interpersonal effectiveness
8. Commitment to learning & personal growth
   Insatiable curiosity; self-awareness & development
9. Commitment to stakeholders
   Customers; team members; the organization; shareholders; the community

The character to lead

10. Personal integrity & ethics
    Personal integrity; ethics; openness
11. Organizational integrity & courage
    Organizational integrity; courage; power
12. Humility, gratitude, & forgiveness
    Humility; gratitude; forgiveness

Appendix C

Qualitative CPA Model (from MAP)


Explicit Feedback
Implicit Feedback
Orienting Response
Triggered
Not Triggered
Received
Processed
Defensiveness
Not Used
Used
Quit/Give Up
Persevereance
Resilience
Optimism Vigor
Humor
Adaptability
Performance
Feedback (IV)
CPA
Feedback Orientation Stages
Performance Outcome (DV)

Appendix D

WLM-BFF Factor and Dimension Breakdown

- **Capacity**
  - Communicate & Influence:
    - Conflict Management
      - Encourages others to voice their needs and concerns in potentially adversarial situations
  - Persevere & Adapt:
    - Energy, Adaptability, and Humor:
      - Can hear bad news and negative feedback without becoming defensive.
- **Commitment**
  - People and Relationships
    - Interpersonal Effectiveness
      - Listens carefully and accurately discerns others' needs and concerns.
    - Collaboration & Teamwork
      - Establishes an environment in which the ideas and suggestions of others are encouraged
  - Learning & Personal Growth
    - Self-Awareness & Development
      - Demonstrates a keen awareness of own motives, behavior, and impact on others.
      - Seeks opportunities to learn about self and elicit feedback from others.
      - Translates developmental feedback and lessons of experience into action.
      - Seeks new challenges and opportunities for personal and professional growth.
- **Character**
  - Personal Integrity & Ethics
    - Personal Integrity
      - Acknowledges and takes responsibility for own mistakes.
    - Humility, Gratitude, & Forgiveness
      - Accepts praise graciously

Appendix E

Worthy Leadership Model - Behavioral Feedback Focus (WLM-BFF)

1. Encourages others to voice their needs and concerns in potentially adversarial situations.
2. Can hear bad news and negative feedback without becoming defensive.
3. Listens carefully and accurately discerns others' needs and concerns.
4. Establishes an environment in which the ideas and suggestions of others are encouraged.
5. Demonstrates a keen awareness of own motives, behavior, and impact on others.
6. Seeks opportunities to learn about self and elicit feedback from others.
7. Translates developmental feedback and lessons of experience into action.
8. Seeks new challenges and opportunities for personal and professional growth.
9. Acknowledges and takes responsibility for own mistakes.
10. Accepts praise graciously.

Appendix F

Feedback Orientation Scale (FOS)

Administer these items using a 5-point scale from *strongly disagree* to *strongly agree*.

Utility
1. Feedback contributes to my success at work.
2. To develop my skills at work, I rely on feedback.
3. Feedback is critical for improving performance.
4. Feedback from supervisors can help me advance in a company.
5. I find that feedback is critical for reaching my goals.

Accountability
1. It is my responsibility to apply feedback to improve my performance.
2. I hold myself accountable to respond to feedback appropriately.
3. I don't feel a sense of closure until I respond to feedback.
4. If my supervisor gives me feedback, it is my responsibility to respond to it.
5. I feel obligated to make changes based on feedback.

Social Awareness
1. I try to be aware of what other people think of me.
2. Using feedback, I am more aware of what people think of me.
3. Feedback helps me manage the impression I make on others.
4. Feedback lets me know how I am perceived by others.
5. I rely on feedback to help me make a good impression.

Feedback Self-Efficacy
1. I feel self-assured when dealing with feedback.
2. Compared to others, I am more competent at handling feedback.
3. I believe that I have the ability to deal with feedback effectively.
4. I feel confident when responding to both positive and negative feedback.
5. I know that I can handle the feedback that I receive.

Appendix G

360˚ (T1) and Pulse (T2) Survey Items

How would you rate this person's overall performance in the past year?
To what extent does this individual demonstrate leadership worth following?
1 = Below average; 5 = Significantly Above Average
1 = To No Extent; 5 = To a Very Great Extent

Appendix H

Pulse Survey (T2) Observation Change Item

To what extent have you observed this individual make positive changes in his/her behavior over the last year?*

*Some participants may receive a low rating on this question because they were previously seen as performing well and not in need of large changes.

1 = To No Extent; 5 = To a Very Great Extent

Appendix I

Study 1 Survey

Welcome to My Survey

Thank you for participating in this survey, A Focus on Feedback at Work.

This survey contains 37 multiple choice questions, and will take between 5 and 10 minutes of your time to complete.

Please note that that you will be asked to agree to an informed consent form on the next page, prior to beginning the survey.

Thank you again for your participation.

Informed Consent

Informed Consent - Online Survey

PRINCIPAL INVESTIGATOR
Aaron Friedman, MS
The University of Texas at Arlington
Department of Psychology
aaron.friedman@mavs.uta.edu

FACULTY ADVISOR
Dr. Nicolette Lopez
Department of Psychology
nlopez@uta.edu

TITLE OF PROJECT
A Focus on Feedback at Work

INTRODUCTION
You are being asked to participate in a research study about your focus on receiving feedback at work. Your participation is voluntary. Refusal to participate or discontinuing your participation at any time will involve no penalty or loss of benefits to which you are otherwise entitled. Please ask questions if there is anything you do not understand.

PURPOSE
The specific purpose of this research study is/are as follows:
-To validate a scale measuring one's focus on feedback
-To understand the relationship between feedback focus and feedback orientation

DURATION
Participation in this study will last approximately 5-10 minutes and can be completed entirely online.

NUMBER OF PARTICIPANTS
The number of anticipated participants in this research study is 120.

PROCEDURES
The procedures, which will involve you as a research participant, include responding to an online survey. First, you as the participant will need to agree to and accept the informed consent page. Once you have agreed, you will respond to the survey, which measures three different components:

1. Demographics, six items including: Age, ethnicity, level of education, industry employed, level of leadership, number of direct report employees
2. Behavioral Feedback Focus Scale, ten items related to seeking and be receptive to feedback at work
3. Feedback Orientation Scale, twenty items related to the ability, or willingness, to receive, process, and use feedback at work

POSSIBLE BENEFITS

The possible benefits of your participation are: You will be contributing to the understanding feedback focus at work.

POSSIBLE RISKS/DISCOMFORTS

There are no perceived risks or discomforts for participating in this research study. Should you experience any discomfort please inform the researcher, you have the right to quit any study procedures at any time at no consequence.

COMPENSATION

The panel that recruited you for this survey will be responsible for your compensation. In this case, the panel has committed to provide 13 cents (US) for a 6 – 9 minute survey, and 25 cents for a 10 – 14 minute survey, all of which can be accumulated and redeemed through Amazon. No additional compensation will be awarded for being part of this research study.

ALTERNATIVE PROCEDURES

You have the option to decline participation in this study by closing your browser.

VOLUNTARY PARTICIPATION

Participation in this research study is voluntary. You have the right to decline participation in any or all study procedures or quit at any time at no consequence.

CONFIDENTIALITY

Every attempt will be made to see that your study results are kept confidential. All data collected from this study will be stored digitally in Dr. Nicolette Lopez' office (LS 308) for at least three (3) years after the end of this research. The results of this study may be published and/or presented at meetings without naming you as a participant. Additional research studies could evolve from the information you have provided, but your information will not be linked to you in anyway; it will be anonymous. Although your rights and privacy will be maintained, the Secretary of the Department of Health and Human Services, the UTA Institutional Review Board (IRB), and personnel particular to this research have access to the study records. Your records will be kept completely confidential according to current legal requirements. They will not be revealed unless required by law, or as noted above. The IRB at UTA has reviewed and approved this study and the information within this consent form. If in the unlikely event it becomes necessary for the Institutional Review Board to review your research records, the University of Texas at Arlington will protect the confidentiality of those records to the extent permitted by law.

CONTACT FOR QUESTIONS

Questions about this research study may be directed to Aaron Friedman at aaron.friedman@mavs.uta.edu or Dr. Nicolette Lopez at nlopez@uta.edu. Any questions you may have about your rights as a research participant or a research related injury may be directed to the Office of Research Administration; Regulatory Services at 817.272.2105 or regulatoryservices@uta.edu.

CONSENT

By clicking "agree and continue" below, you confirm that you are 18 years of age or older and have read or had this document read to you. You have been informed about this study's purpose, procedures, possible benefits and risks. You have been given the opportunity to ask questions before you agree, and you have been told that you can ask other questions at any time by emailing the primary researcher. You voluntarily agree to participate in this study. By clicking "agree and continue" below, you are not waiving any of your legal rights. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue participation at any time without penalty or loss of benefits, to which you are otherwise entitled. If you agree, please select the option below before continuing with the survey. Should you disagree, please close your internet browser immediately by clicking the X at the top of your browser (top right for PCs and top left for Macs), and do not select the option below or continue with the survey.

I agree and choose to continue by clicking the bubble below.

Agree and continue

WLM-BFF

Please indicate the extent to which you demonstrate the following behaviors at work:

I encourage others to voice their needs and concerns in potentially adversarial situations.

○ To No Extent ○ To a Little Extent ○ To Some Extent ○ To a Great Extent ○ To a Very Great Extent

I can hear bad news and negative feedback without becoming defensive.

○ To No Extent ○ To a Little Extent ○ To Some Extent ○ To a Great Extent ○ To a Very Great Extent

I listen carefully and accurately discern others' needs and concerns.

○ To No Extent ○ To a Little Extent ○ To Some Extent ○ To a Great Extent ○ To a Very Great Extent

I establish an environment in which the ideas and suggestions of others are encouraged.

○ To No Extent ○ To a Little Extent ○ To Some Extent ○ To a Great Extent ○ To a Very Great Extent

I demonstrate a keen awareness of my own motives, behavior, and impact on others.

○ To No Extent ○ To a Little Extent ○ To Some Extent ○ To a Great Extent ○ To a Very Great Extent

I seek opportunities to learn about myself and elicit feedback from others.

◯ To No Extent ◯ To a Little Extent ◯ To Some Extent ◯ To a Great Extent ◯ To a Very Great Extent

I translate developmental feedback and lessons of experience into action.

◯ To No Extent ◯ To a Little Extent ◯ To Some Extent ◯ To a Great Extent ◯ To a Very Great Extent

I seek new challenges and opportunities for personal and professional growth.

◯ To No Extent ◯ To a Little Extent ◯ To Some Extent ◯ To a Great Extent ◯ To a Very Great Extent

I acknowledge and take responsibility for my own mistakes.

◯ To No Extent ◯ To a Little Extent ◯ To Some Extent ◯ To a Great Extent ◯ To a Very Great Extent

I accept praise graciously.

◯ To No Extent ◯ To a Little Extent ◯ To Some Extent ◯ To a Great Extent ◯ To a Very Great Extent

FOS

Please indicate the extent to which you agree with the following statements:

Feedback contributes to my success at work.

○ Strongly Disagree ○ Disagree ○ Neither Agree nor Disagree ○ Agree ○ Strongly Agree

To develop my skills at work, I rely on feedback.

○ Strongly Disagree ○ Disagree ○ Neither Agree nor Disagree ○ Agree ○ Strongly Agree

Feedback is critical for improving performance.

○ Strongly Disagree ○ Disagree ○ Neither Agree nor Disagree ○ Agree ○ Strongly Agree

Feedback from supervisors can help me advance in a company.

○ Strongly Disagree ○ Disagree ○ Neither Agree nor Disagree ○ Agree ○ Strongly Agree

I find that feedback is critical for reaching my goals.

○ Strongly Disagree ○ Disagree ○ Neither Agree nor Disagree ○ Agree ○ Strongly Agree

It is my responsibility to apply feedback to improve my performance.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I hold myself accountable to respond to feedback appropriately.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I don't feel a sense of closure until I respond to feedback.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

If my supervisor gives me feedback, it is my responsibility to respond to it.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I feel obligated to make changes based on feedback.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I try to be aware of what other people think of me.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

Using feedback, I am more aware of what people think of me.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

Feedback helps me manage the impression I make on others.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

Feedback lets me know how I am perceived by others.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I rely on feedback to help me make a good impression.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I feel self-assured when dealing with feedback.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

Compared to others, I am more competent at handling feedback.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I believe that I have the ability to deal with feedback effectively.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I feel confident when responding to both positive and negative feedback.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

I know that I can handle the feedback that I receive.

◯ Strongly Disagree ◯ Disagree ◯ Neither Agree nor Disagree ◯ Agree ◯ Strongly Agree

Appendix J

PCA Scree Plot of the WLM-BFF


Scree Plot
Eigenvalue
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10
Component Number

References

Boyce, A. S., Corbet, C. E., & Adler, S. (2013). Simulations in the selection context: considerations, challenges, and opportunities. In Fetzer, M., & Tuzinski, K. (2013). *Simulations for personnel selection.* New York: Springer Science + Business Media.

Braddy, P. W., Sturm, R. E., Atwater, L. E., Smither, J. W., & Fleenor, J. W. (2013). Validating the feedback orientation scale in a leadership development context. *Group & Organization Management, 38*(6), 690-716.

Carlson, K. D., & Herdman, A. O. (2012). Understanding the impact of convergent validity on research results. *Organizational Research Methods*, *15*(1), 17-32.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Rev. ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Dahling, J. J., Chau, S. L., & O'Malley, A. (2012). Correlates and consequences of feedback orientation in organizations. *Journal of Management, 38,* 531-546.

Dolbier, C. L., Webster, J. A., McCalister, K. T., Mallon, M. W., & Steinhardt, M. A. (2005). Reliability and validity of a single-item measure of job satisfaction. *American Journal of Health Promotion, 19*(3), 194-198.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Gluck, M., Mercado, E., & Myers, C. (2011). *Learning and memory: From brain to behavior* (2nd ed.). New York: Worth Publishers.

Gough, H. & Bradley, P. (2005). *CPI 260 Manual.* Mountain View, CA: CPP, Inc.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26,* 449-510.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology, 64,* 349-371.

Jordan, J. S., & Turner, B. A. (2008). The feasibility of single-item measures for organizational justice. *Measurement in Physical Education and Exercise Science, 12*(4), 237-257.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*(4), 765-780.

Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology, 89*(3), 542-552.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36.

Lawrence, H. V. & Wiswell, A. K. (1993). Using the work group as a laboratory for learning: Increasing leadership and team effectiveness through feedback. *Human Resource Development Quarterly, 4*(2), 135-148.

Linderbaum, B. A., & Levy, P. E. (2010). The development and validation of the feedback orientation scale (FOS). *Journal of Management, 36,* 1372-1405.

London, M., & Smither, J. W. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review, 12,* 81-100.

Manoogian, S. (2006). *CPI 260 Coaching Report for Leaders: Advanced Guide for Interpretation.* Mountain View, CA: CPP, Inc.

Moriarty, K. O. (2003). *Are multiple rater perspectives better? A test of two assumptions.* Dissertation Abstracts International, 64, 991.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Oh, I., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology, 96*(4), 762-773.

Riggio, R. (2000). *Introduction to industrial/organizational psychology* (3rd ed.). Upper Saddle River, NJ: Prentice Hall, Inc.

Roth, P., Bobko, P., McFarland, L., & Buster, M. (2008). Work sample tests in personnel selection: A meta-analysis of Black-White differences in overall and exercise scores. *Personnel Psychology, 61*(3), 637-662

Society for Consulting Psychology (SCP; 2015). http://www.apadivisions.org/division-13/

Steelman, L. A., Levy, P. E., & Snell, A. F. (2004). The feedback environment scale: Construct definition, measurement, and validation. *Educational and Psychological Measurement, 64*, 165-184.

Tabachnick, B. & Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.

Thompson, A. D. (1987). A formative evaluation of an individualized coaching program for business managers and professionals (Doctoral dissertation, University of Minnesota, 1987). *Dissertation Abstracts International, 47*(12), 4339A.

Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology, 39,* 212-222.

Whetzel, D. L., McDaniel, M. A., & Nguyen, N. T. (2008). Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance, 21*(3), 291-309

Yukl, G., & Mahsud, R. (2010). Why flexible and adaptive leadership is essential. *Consulting Psychology Journal: Practice and Research, 62*(2), 81-93.

Worthyleadership.com (2015). http://worthyleadership.com. Retrieved May 1, 2015.

Biographical Information

Aaron D. Friedman graduated from the University of Oklahoma in 2007 with a BA in Psychology. Aaron received his M.S. in Industrial & Organizational Psychology in 2013 and Ph.D. in Experimental Psychology in 2015 from the University of Texas at Arlington. Aaron's primary area of interest is in Executive Assessment, Executive Coaching, and Consulting Psychology.