

YARBOROUGH, J. PRESTON, Ph.D. The Development and Validation of the Leadership Versatility Index for Students (LVI-S). (2011)
Directed by Dr. James M. Benshoff and Dr. Craig S. Cashwell pp. 350

According to Bass' (1990) summary of fifty years of research and nearly thirty dichotomy-based theories, leaders influence people through autocratic use of power (task-oriented) or through democratic use of power (people-oriented). Each style produces unique tensions and tradeoffs, but versatile leaders can incorporate strategies from both sides of the dichotomy, depending on situational needs (Kaplan, 1996). Versatile leaders avoid overusing strengths to the point of weakness—a frequently overlooked leadership flaw (Kaplan & Kaiser, 2006). The versatile leader concept shares much with synergistic supervision, a student affairs supervision model (Winston & Creamer, 1997; 1998). Synergistic supervisors blend strengths from autocratic and democratic approaches, creating synergistic relationships with those they lead (Winston & Creamer, 1997; 1998). Synergy and versatility may be considered different sides of the same coin.

Until the Leadership Versatility Index—Student (LVI-S), no quantitative, multi-rater measure of leadership versatility was available for campus leaders. The LVI-S was derived from the executive-focused Leadership Versatility Index[®] (Kaplan and Kaiser, 2006). Participants were recruited from departments of housing and residence life across seven institutions in the Southeastern United States, including staff from small private colleges through large public universities. Resident Advisor supervisees ($n = 262$) rated leadership characteristics of their Hall Directors ($n = 52$); the study averaged 4.9 raters-per-leader. Convergent validity was tested using the Student Leadership Practices

Inventory[®] (SLPI) (Kouzes & Posner, 2003); predictive validity was tested through a global effectiveness measure derived from Tsui's (1984) effectiveness research.

LVI-S scale alphas exceeded .80 and scales offered compelling evidence of convergent and predictive validity. A strong predictive relationship was found between versatility and effectiveness ($R = .60$, $\text{Adj. } R = .31$, $F = 7.72$, $p < .01$). Results validated the LVI-S for use in residence life settings and validated behavioral aspects of synergistic supervision. Applications for the LVI-S were discussed as well as avenues for future research.

THE DEVELOPMENT AND VALIDATION OF THE LEADERSHIP VERSATILITY
INDEX FOR STUDENTS (LVI-S)

by

J. Preston Yarborough

A Dissertation Submitted to
The Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2011

Approved by

Dr. James M. Benshoff
Committee Co-Chair

Dr. Craig S. Cashwell
Committee Co-Chair

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of
The Graduate School at The University of North Carolina at Greensboro.

| | |
|--------------------|-----------------------------|
| Committee Co-chair | <hr/> Dr. James M. Benshoff |
| Committee Co-chair | <hr/> Dr. Craig S. Cashwell |
| Committee Members | <hr/> Dr. Terry Ackerman |
| | <hr/> Dr. Kevin B. Lowe |

12/6/2010
Date of Acceptance by Committee

12/6/2010
Date of Final Oral Examination

ACKNOWLEDGEMENTS

This project was the result of an extraordinary community effort. First, I would like to thank my committee members. Leadership is not a common topic of study for counselor educators. Their patience and persistence made an enormous difference in the outcomes of this project. Second, the institutions participating were extraordinarily generous with their time, information, and assistance. The complexity and magnitude of this study caused it to run longer than anticipated, but I hope the results prove to be worth the wait. Their patience is greatly appreciated.

The scholarly contributions and cooperation of Kaplan DeVries, Inc. (particularly Rob Kaiser, Darren Overfield, and Bart Craig) are greatly appreciated. This project could not have been completed without your ongoing support. Jim Kouzes and Wiley Publishing are greatly appreciated as well, for allowing the SLPI to be used in this study.

The greater High Point and Greensboro restaurants that supported these efforts should also be noted. The staff of Moe's Southwestern Grill, McAlister's Deli, and Panera all have my continued patronage and gratitude. Thanks for the many bagels, teas, chili bowls, and burritos that sustained me through six years of study.

I want to thank the friends that provided support through the many trials of this process. Chris Ward, Cathy Hamilton (the staff of the Office of Leadership and Service-Learning), Marin Burton, Paul Lessard, David Sanford, Rick Bender, Gerald Hutchinson, Glenn Newsome, and many others have been appreciated in ways I can't begin to fathom.

Finally, a special thanks to my family. Thanks Mom and Scott for reminding me that Yarboroughs persevere. We know how proud Dad would be and that's something we can share. It's a celebration for him that we've made it this far.

The only Daddy my kids have known has been one writing a dissertation. Their support and prayers, their understanding when I had to leave on weekends to write; although it pains me in many ways, their unconditional love will always be deeply felt. They have been generous, even at an age when they shouldn't have to be.

Mendy, you know more than I do the sacrifices our family has made. Thank you for understanding and for sticking with me. We never knew it would be this hard. I love you with all my heart and I want you to know this was completed for all of us, not only for me. You accept me even when you don't understand me—and that's a love that's tough to describe. Thanks for everything.

TABLE OF CONTENTS

| | Page |
|---|------|
| LIST OF TABLES | xii |
| LIST OF FIGURES | xiv |
| CHAPTER | |
| I. INTRODUCTION | 1 |
| The Versatile Leader Model and the Leadership Versatility | |
| Index | 6 |
| Versatility and Lopsidedness | 7 |
| Conceptual Structure of the LVI | 8 |
| Too Little Too Much response scale | 8 |
| A duality-based structure | 9 |
| The Forceful / Enabling duality | 10 |
| Takes Charge versus Empowers | |
| Others sub-dimensions | 10 |
| Declares versus Listens sub- | |
| dimension | 11 |
| Pushes versus Supports sub- | |
| dimension | 11 |
| The Strategic / Operational duality | 11 |
| Summary of conceptual structure | 12 |
| Empirical Support for the LVI Hierarchical Structural | |
| Model | 13 |
| The Value of the LVI as a Development Tool | 15 |
| Developing Future Leaders | 16 |
| Statement of the Problem | 26 |
| Purpose of the Study | 28 |
| Research Questions | 29 |
| Need for the Study | 30 |
| Key Terms and Definitions | 31 |
| Brief Overview | 32 |
| II. LITERATURE REVIEW | 34 |
| The Contemporary Leadership Development Context | 34 |
| Are Leaders Born or Made? | 37 |
| Progression of leadership theories | 37 |
| Style and behavioral theories | 37 |

| | |
|--|----|
| Contingency theories..... | 38 |
| New schools of study..... | 40 |
| Research advancements..... | 41 |
| Advances in methods, analyses, and instrumentation..... | 41 |
| Developing Leadership in Corporate Environments..... | 43 |
| Competency Models..... | 43 |
| Derailment behaviors and metacompetencies..... | 44 |
| Problems with competency models..... | 45 |
| Measuring Competency Models: Multi-rater Response Formats..... | 45 |
| The frequency response format..... | 46 |
| The evaluation response format..... | 46 |
| Blind spots in measurement formats..... | 47 |
| Blind spots in the frequency format..... | 47 |
| Blind spots in the evaluative format..... | 48 |
| Implications of blind spots..... | 48 |
| Summary of measurement..... | 49 |
| The Versatile Leader Model..... | 50 |
| Leadership Versatility..... | 50 |
| Conceptual structure of the Leadership Versatility Index..... | 52 |
| Too Little/Too Much response scale..... | 52 |
| Duality-based structure..... | 53 |
| Forceful / Enabling duality..... | 55 |
| Perspective one: Locus of power..... | 56 |
| Perspective two: Decision- making..... | 56 |
| Perspective three: Orientation..... | 57 |
| Strategic / Operational duality..... | 58 |
| Psychometric Properties of the Leadership Versatility Index..... | 59 |
| The Forceful / Enabling factorial structure..... | 59 |
| Reliability of the LVI..... | 60 |
| Validity of the LVI..... | 61 |
| Value of LVI as a Development Tool..... | 62 |
| “The Future That Has Already Happened”..... | 64 |
| Shifting Leadership Paradigms..... | 66 |
| Development of Leadership in College and University Contexts..... | 67 |
| How Do Leaders Develop in College and University Settings?..... | 69 |
| Concerns with Student Affairs Leadership Literature..... | 70 |
| Methodological Concerns..... | 70 |
| Critical Analysis of Self-report Validity and Student Affairs Research Methods..... | 73 |

| | |
|---|-----|
| A meta-analysis of self-report data | 75 |
| Problems with self-report leadership data..... | 77 |
| Summary for conditions of validity and limitations of self-reported data | 80 |
| Comparisons of Self-reported Data to Observer-reported Data | 81 |
| Multi-rater meta-analysis incorporating peer and supervisor ratings. | 81 |
| Multi-rater meta-analysis incorporating peer, superior, and supervisee ratings. | 83 |
| Implications of directional ratings. | 85 |
| The self-ratings of the unskilled and unaware | 85 |
| Summary comparing self-reported data to observer-reported data | 90 |
| Levels of Analysis..... | 91 |
| The ecological fallacy..... | 92 |
| Self-awareness and multi-rater methodology. | 94 |
| Summary of levels of analysis | 99 |
| Development of Leadership in Housing and Residence Life Contexts..... | 100 |
| Student Leadership Context: Housing and Residence Life | 102 |
| Paraprofessionals in housing and residence-life. | 102 |
| The Resident Advisor role. | 103 |
| The Hall Director role..... | 104 |
| Leadership Styles and Leadership Development in Residence Hall Settings | 106 |
| General Observations About Paraprofessional Residence Hall Leadership and Governance | 106 |
| Interpersonal Characteristics and Skills of Paraprofessionals..... | 107 |
| Paraprofessional Role Clarity | 108 |
| Paraprofessional Supervisory Styles..... | 109 |
| The Authoritarian Supervisor. | 110 |
| The Laissez-faire Supervisor. | 111 |
| The RA's RA. | 111 |
| The Synergistic Supervisor. | 112 |
| Development and Supervision of Student Affairs Professionals..... | 113 |
| Problematic Approaches to Supervision..... | 114 |
| Systematic Development and Supervision of Student Affairs Professionals | 116 |
| High quality supervision..... | 117 |

| | |
|--|-----|
| Measuring the Development and Supervision of Student Affairs Professionals | 120 |
| Summary of Paraprofessional and Professional Development, Supervision, and Leadership | 121 |
| Contemporary Leadership Models in Residence Hall Leadership: Theory and Measurement | 123 |
| Situational Leadership | 123 |
| Transformational Leadership | 124 |
| Transformational leadership according to Bass and Avolio. | 125 |
| Assessing leadership using Bass and Avolio’s transformational leadership model. | 126 |
| Housing and residence life studies using the Multifactor Leadership Questionnaire..... | 128 |
| Transformational factors, RA satisfaction, motivation, and HD effectiveness. | 128 |
| Transformational factors and achievement styles..... | 130 |
| Transformational factors, gender differences and RA satisfaction..... | 132 |
| The Five Practices Model | 133 |
| Measurement of Kouzes and Posner’s model..... | 134 |
| Five Practices and RA leadership and effectiveness. | 135 |
| Effectiveness as a Criterion Variable in Residence Hall Multi-rater Studies..... | 138 |
| Questionnaire-based effectiveness measures and residence halls. | 138 |
| Meta-analysis of Multifactor Leadership Questionnaire research. | 139 |
| Distinct rater groups in HRL research. | 140 |
| Summary | 141 |

III. METHODOLOGY 145

| | |
|--|-----|
| Research Questions and Hypotheses | 145 |
| Participants..... | 147 |
| Instrumentation | 148 |
| Leadership Versatility Index-Student (LVI-S) | 148 |
| The Leadership Practices Inventory Student Version (SLPI) | 153 |
| Reputational Effectiveness Scale..... | 154 |
| Demographic Questionnaire | 155 |

| | |
|---|---------|
| Procedures..... | 157 |
| Modifying the LVI into the Leadership Versatility Index—Student (LVI-S)..... | 157 |
| Establishing the Scope of the LVI-S..... | 157 |
| Item Modifications for the LVI-S..... | 158 |
| Development of LVI-S Assessment Platform..... | 163 |
| Commercial LVI assessment platform..... | 163 |
| LVI-S administration platform..... | 164 |
| Administration of the Study..... | 165 |
| Data Analysis..... | 167 |
| IV. RESULTS | 170 |
| Sample Characteristics..... | 170 |
| Hall Directors..... | 170 |
| Resident Advisors..... | 174 |
| Instrument Psychometrics..... | 176 |
| Research Hypotheses | 183 |
| Research Question 1 | 183 |
| Principal Components Analysis: Forceful Dimension..... | 185 |
| Principal Components Analysis: Enabling Dimension | 188 |
| Inverse relationship of the Forceful and Enabling dimensions..... | 190 |
| Research Question 2 | 191 |
| Research Question 3 | 193 |
| Research Question 4 | 195 |
| Research Question 5 | 198 |
| Forceful dimension analysis..... | 199 |
| Enabling dimension analysis..... | 200 |
| Direction of rating discrepancies..... | 202 |
| Enabling dimension agreement groups and correlates of effectiveness..... | 208 |
| Research Question 6 | 211 |
| V. CONCLUSIONS..... | 220 |
| Overview..... | 220 |
| Discussion..... | 224 |
| Estimates of Reliability..... | 224 |
| Internal consistency..... | 225 |
| Inter-rater reliability..... | 226 |

| | |
|--|-----|
| Measuring Leadership Versatility in a Housing and Residence | |
| Life Context..... | 228 |
| Research Question 1 | 229 |
| Research Question 2 | 234 |
| Research Question 3 | 240 |
| Research Question 4 | 244 |
| Research Question 5 | 247 |
| Research Question 6 | 251 |
| Limitations | 254 |
| Implications for College / University Counseling | 255 |
| Suggestions for Future Research | 258 |
| Conclusion | 259 |
| REFERENCES | 263 |
| APPENDIX A. HALL DIRECTOR RESPONSE FORM USED FOR MAIN STUDY | 289 |
| APPENDIX B. RESIDENT ADVISOR RESPONSE FORM USED FOR MAIN STUDY..... | 296 |
| APPENDIX C. LVI-S FIELD TEST ITEMS FOR PILOT STUDY | 303 |
| APPENDIX D. ITEMS SUBMITTED FOR EXPERT REVIEW..... | 309 |
| APPENDIX E. FIRST REVISED SET OF ITEMS | 311 |
| APPENDIX F. INITIAL SET OF PROPOSED ITEMS | 313 |
| APPENDIX G. PERMISSION FOR LEADERSHIP PRACTICES INVENTORY | 315 |
| APPENDIX H. PERMISSION FOR LEADERSHIP VERSATILITY INDEX | 316 |
| APPENDIX I. INFORMED CONSENT, FOCUS GROUP PARTICIPANT..... | 317 |
| APPENDIX J. INFORMED CONSENT, PILOT STUDY HALL DIRECTOR | 320 |
| APPENDIX K. INFORMED CONSENT, EXPERT REVIEW..... | 323 |
| APPENDIX L. PILOT STUDY SCRIPT | 326 |
| APPENDIX M. PILOT STUDY RECRUITMENT E-MAIL | 328 |

| | |
|---|-----|
| APPENDIX N. MAIN STUDY HALL DIRECTOR RECRUITMENT | 329 |
| APPENDIX O. MAIN STUDY FOLLOW-UP HALL DIRECTOR RECRUITMENT | 330 |
| APPENDIX P. MAIN STUDY RA RECRUITMENT LETTER—HALL DIRECTOR..... | 331 |
| APPENDIX Q. MAIN STUDY FOLLOW-UP RA RECRUITMENT LETTER..... | 332 |
| APPENDIX R: HALL DIRECTOR SUPERVISOR RECRUITMENT LETTER..... | 333 |
| APPENDIX S. INSTITUTIONAL LETTER OF SUPPORT EXAMPLE..... | 334 |
| APPENDIX T. MAIN STUDY INFORMED CONSENT, HALL DIRECTOR | 335 |
| APPENDIX U. INFORMED CONSENT, RESIDENT ADVISOR | 338 |
| APPENDIX V. INFORMED CONSENT SUPERVISOR OF HALL DIRECTOR..... | 341 |
| APPENDIX W. ADMINISTRATIVE PROCESSES AND PROCEDURES | 344 |
| APPENDIX X. EXPLANATION OF JOINT-STANDING SCORES | 345 |
| APPENDIX Y. LVI-S SUB-DIMENSION INTER-ITEM CORRELATIONS..... | 349 |
| APPENDIX Z. LVI-S CONCEPTUAL STRUCTURE DIAGRAM | 350 |

LIST OF TABLES

| | Page |
|---|------|
| Table 1. Research Questions..... | 168 |
| Table 2. Hall Director Sample Characteristics | 171 |
| Table 3. Distribution of HD Sample Across Institutional Type/Residence Type..... | 172 |
| Table 4. Number of RAs Supervised Versus Number of RA Participants per Institution Type | 172 |
| Table 5. HD Participation Rate by Institutional Type | 173 |
| Table 6. Race/Ethnicity of Hall Directors by Institution and Residence Hall Type..... | 173 |
| Table 7. Gender of Hall Directors by Institution and Residence Hall Type..... | 173 |
| Table 8. RA Sample Characteristics | 174 |
| Table 9. Distribution of RA Sample Across Institutional Type/Residence Type..... | 175 |
| Table 10. RA Participation Rate by Institutional Type..... | 175 |
| Table 11. Race/Ethnicity of RAs by Institution..... | 176 |
| Table 12. Gender of Resident Advisors by Institution | 176 |
| Table 13. <i>ICC</i> Analyses Sub-sample: Demographic Characteristics and Descriptive Statistics | 181 |
| Table 14. <i>ICC</i> Sub-sample Crosstab Analysis: RA Year by RA Experience (in years)..... | 181 |
| Table 15. Internal Consistency and Inter-rater Reliability..... | 182 |
| Table 16. Component Matrix (with Varimax rotation): Forceful Dimension | 184 |
| Table 17. Component Matrix (with Varimax rotation): Enabling Dimension..... | 185 |
| Table 18. Variance Explained: HD LVI-S Sub-dimension Scores and Composite Effectiveness Score | 191 |

| | |
|--|-----|
| Table 19. Coefficients ^a of LVI-S Sub-dimension Scores and Composite Effectiveness Score | 192 |
| Table 20. Variance Explained: HD (Self-report) Sub-dim Scores and RA (Observer-report) Sub-dim Scores | 193 |
| Table 21. Pearson's <i>r</i> Correlation Matrix for LVIS Sub-dimensions, LPI Scales, and Reputational Effectiveness Scale..... | 194 |
| Table 22. Means Comparison of HD and RA LVI-S Sub-dimension Scores..... | 196 |
| Table 23. Dimension-level Mean Difference Scores and Boundary Parameters for Rater Agreement Categories..... | 197 |
| Table 24. Crosstab Analysis: Rater Assignment Categories and Chi Square Test..... | 197 |
| Table 25. Forceful Dimension Mean Comparisons: Agreement Categories by Composite Effectiveness (Eff_Comp) Score | 199 |
| Table 26. Enables Dimension Mean Comparisons: Agreement Categories by Composite Effectiveness (Eff_Comp) Score | 201 |
| Table 27. Enables Dimension Mean Comparisons: Agreement Categories and Correlates of Effectiveness by Rater Source..... | 209 |
| Table 28. Enables Sub-dimension Mean Comparisons: Agreement Categories and Correlates of Effectiveness by Rater Source | 210 |
| Table 29. Additive and Versatility Model Comparison (Split Sample) Predicting Effectiveness | 215 |
| Table 30. Variance Explained: SLPI Scales | 215 |

LIST OF FIGURES

| | Page |
|---|------|
| Figure 1. Too Little Too Much Response Format | 150 |
| Figure 2. Calculation of Joint-standing Scores | 152 |
| Figure 3. Instruments, Scales, and Subscales | 156 |
| Figure 4. Forceful Dimension Scree Plot..... | 187 |
| Figure 5. Enabling Dimension Scree Plot..... | 189 |
| Figure 6. Rater Agreement Categories and Effectiveness | 204 |
| Figure 7. Under-estimators and Effectiveness | 205 |
| Figure 8. Over-estimators and Effectiveness | 206 |
| Figure 9. Self-Awareness and Effectiveness | 207 |
| Figure 10. Distribution of Forceful and Enabling Versatility Plots..... | 218 |

CHAPTER I

INTRODUCTION

Great leadership benefits the collective, whether at the national, community, or organizational level (Hogan & Kaiser, 2005). The consequences of poor leadership can be devastating (Hogan & Kaiser, 2005). According to noted scholars Astin and Astin (2002), the quality of leadership in the United States is in decline. As evidence of this degradation, Astin and Astin (2002) cited community concerns such as eroding social trust, reduced civic engagement, and increasingly inequitable distribution of wealth. To further complicate matters, the baby boomer generation is quickly approaching retirement. In the coming decade vast numbers of experienced leaders are expected to retire or to step down from full-time employment; this shift will create a leadership vacuum (Teegarden, 2004; Weik, 2005). By necessity, these vacated positions will be filled by a younger, but less populous generation of workers (Drucker, 1998; Rappaport, Bancroft, & Okum, 2003). As organizations adapt, they will become “responsibility-based organizations in which every member must act as a responsible decision maker. All members, in other words, have to see themselves as ‘executives’” (Drucker, 1998, p. 126). Therefore, workers at all organizational levels must develop competent leadership skills. This dilemma prompts several critical questions: 1) Is Drucker’s speculation realistic? Does everyone have the potential to learn leadership? 2) If Drucker’s proposal were implemented, when should we begin the leadership development process?

In the late 19th and early 20th century, leadership was considered a product of inborn traits possessed only by exceptional individuals (Bass, 1990). Subsequent to trait theory, researchers determined that leadership was also a function of behaviors (Bales, 1954; Kahn & Katz, 1953; Stogdill & Coons, 1957) and of situational variables (Fiedler, 1967; House, 1971; Vroom & Yetton, 1973). Behavioral and situational perspectives represented an important theoretical shift: Leadership could be learned. If leadership can be learned, then leaders can be developed (McCauley & Van Velsor, 2004).

To best teach leadership, leader development programs needed to identify and convey the essential ingredients for great leadership, but this task proved challenging. Behavioral theories failed to identify a set of behaviors that consistently produced effective outcomes (House & Aditya, 1997). Evidence was inconclusive or contradictory (Bass, 1990; House & Aditya, 1997). Behavioral theories yielded to contingency theories (House & Aditya, 1997). Contingency theories demonstrated that situational variables moderated outcomes: A behavior that was beneficial in one circumstance could be detrimental in another (Fiedler, 1967; House, 1971; Vroom & Yetton, 1973). Contingency theories generated empirical support, but this support was debated (House & Aditya, 1997). It also appeared that for each question contingency theories answered, several more complex questions emerged (Antonakis, Cianciolo, & Sternberg, 2004). In the 1980's contingency theories faded from prominence due to their complexity and an inability provide specific, practical instruction (Antonakis, et al., 2004).

But findings from contingency theory helped to transform the study of leadership. According to Antonakis, et al, (2004) new schools of study were inspired by contingency

theory: relational leadership (Graen & Uhl-Bien, 1995), information-processing (Lord, Foti, & De Vader, 1984), and transformational theory (Bass, 1985a). These schools isolated elements of contingency theory and examined each in greater depth (Antonakis, et al., 2004). Lowe and Gardner (2000) noted that along with this more granular approach, increasingly sophisticated assessment protocols emerged. Mabe and West (1982) and Podsakoff and Organ (1986) demonstrated limitations and confounds of self-reported data. Studies showed that multi-rater data exhibited greater reliability and validity than self-reported data (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988).

Since the early 1980's, multi-rater instrumentation has proliferated and the psychometric quality of these instruments improved as well (Leslie & Fleenor, 1998). Multi-rater instruments are frequently referred to as 360-degree assessments. This moniker is appropriate in corporate settings because 360-degree assessments measure perspectives from superiors, bosses, peers, and subordinates—a full circle approach to leadership measurement. In multi-rater assessment, multiple raters are asked to provide feedback on their perceptions of the focal leader's behaviors. Feedback is captured with a form that has a common set of questions for each rater. Ratings are anonymous to encourage maximum candor on the part of the raters (Eichinger & Lombardo, 2003). Scores from each rater are compiled and organized into a feedback report for the focal leader. Ideally, this leader reviews the feedback report with an individual skilled in facilitating assessment interpretation and feedback (Leslie & Fleenor, 1998).

As the quality of results improved, researchers incorporated better methodologies and used more effective analytical procedures. The resultant literature provided better

insight into the characteristics, behaviors, and situational variables of effective leadership (Kroeck, Lowe, & Brown, 2004). These advances were reflected in the way leadership was measured and taught in corporate leadership programs (Charan, Drotter, & Noel, 2001).

Leadership competencies, or measurable characteristics related to work success (Boyzatis, 1982; Lombardo & Eichinger, 2000), gained prominence in leadership development programs (Lombardo & Eichinger, 2000). Organizations identified competencies believed to be critical for success in a work-role; they used competency-based training and assessment to enhance employee performance in their current roles as well as to prepare them for future advancement (Lombardo & Eichinger, 2000; Silzer, 2006). Competencies have also been used to inform employee selection (Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski, 2001).

Some competency models included derailment behaviors, or behaviors that could impede successful performance (Lombardo & Eichinger, 2000; McCall & Lombardo, 1983). Derailment behaviors can put a career into peril (Lombardo & Eichinger, 2000). Most derailment behaviors resulted from inappropriately used strengths (Lombardo & Eichinger, 2000; McCall & Lombardo, 1983). In other words, many of the competencies that make leaders effective also lie at the root of their derailment. For this reason, Hollenbeck and McCall (2006) stated, “strengths can become weaknesses (and therefore every competence is also an incompetence) ... effectiveness depends on how various *combinations* of strengths are used and that different strengths and weaknesses come into focus at different times and in different jobs” (pp. 399-400, italics in original). One must

not only learn a competency, but also learn when to use it effectively. This ability has been referred to as behavioral flexibility (Zaccaro, Foti, & Kenny), adaptability (Briscoe & Hall, 1999), behavioral complexity (Denison, Hooijberg, & Quinn, 1995), integrative thinking (Martin, 2007), leadership agility (Lombardo & Eichinger, 2000), and leadership versatility (Kaplan & Kaiser, 2006).

Most leadership instrumentation is limited in its ability to measure versatility (Kaiser & Kaplan, 2005). Response scales contribute to this problem (Kaiser & Kaplan, 2005). Most multi-rater instruments utilize continuous response scales similar to the one Rensis Likert (1932) used to measure attitudes (Kaiser & Kaplan, 2005; Leslie & Fleenor, 1998). A continuous scale provides response options where a number value is assigned to each response, such as 5 = *Always*, 3 = *Sometimes*, and 1 = *Never*. Number values are sequential and generally begin at 0 or 1 and proceed upward. Scores are calculated by adding or averaging across items in a scale (Leslie & Fleenor, 1998). Typically these response scales are configured into *frequency* scales, which measure behavioral frequency, or *effectiveness* scales which measure behavioral effectiveness (Kaiser & Kaplan, 2005; Leslie & Fleenor, 1998). Both frequency and effectiveness formats are affected by a blind spot. They are unable to provide feedback on the overuse of strengths (Kaiser & Kaplan, 2005).

Frequency response scales are good for unearthing deficient behavior, but they are less effective at the high end of the register. A low score means a leader is not performing a skill or ability. High scores, however could mean the skill is used frequently (with good effect) or used too much (with detrimental effect) (Kaiser & Kaplan, 2005). Frequency

response scales ask the rater to describe performance, not judge it (Kaiser & Kaplan, 2005). Therefore “high” ratings on a frequency scale could mask problematic aspects of overused behaviors (Kaiser & Kaplan, 2005).

Evaluative ratings become unclear at the low end of the scale (Kaiser & Kaplan, 2005). When evaluative ratings are high, the recipient knows to continue doing what she has been doing. But if a leader is rated low (e.g., “adequate” or “poor”) on certain leadership abilities, there is little to indicate why the low rating was given. Improvement may be needed, but it is unclear whether it is because of a deficient ability or due to an overused strength (Kaiser & Kaplan, 2005).

The Versatile Leader Model and the Leadership Versatility Index

In the early 1990’s, Kaplan recognized many executives received high 360-degree instrument ratings (high frequencies of leadership behaviors), but co-workers’ narrative comments referenced problems not apparent in the quantitative feedback (Kaplan, 1996). Kaplan (1996) hypothesized that excessive use of strengths could be just as problematic as under-used strengths. The versatile leader model and the Leadership Versatility Index (LVI) were developed to address this shortcoming (Kaplan & Kaiser, 2006). Studies with the LVI showed more than 80% of executive leaders tended to overuse a narrow range of strengths; furthermore, the degree to which they overused strengths predicted lower effectiveness ratings (Kaplan & Kaiser, 2003a).

The root of the word versatility meant “to turn around” or “to pivot” (Kaplan, 1996, p.1). Versatility is the ability to pivot from approach to another depending on the needs of a situation (Kaplan, 1996). Kaplan and Kaiser (2006) observed that versatility

functions across two broad domains: (1) how someone leads and (2) what someone leads (Kaplan & Kaiser, 2006). Interpersonal influence, or *how* leaders lead, addressed the way leaders interact with others (Kaplan & Kaiser, 2006). Sometimes leaders are forceful, imposing their will; but at other times they enable others to chart a course of action (Kaplan, 1996). *What* leaders lead, or the initiatives they champion, is another aspect of versatility (Kaplan & Kaiser, 2006). Leaders may emphasize innovative initiatives that encourage growth or they may emphasize day-to-day operational efficiency. Conflicting interests across these domains, and the tensions and tradeoffs therein, are elemental to the art of leadership (Kaplan & Kaiser, 2006).

Versatility and Lopsidedness

When presented with two opposing approaches, people generally place greater value on the side in which they have greatest familiarity or historical success; conversely, they typically overlook the value of its opposite (Kaplan & Kaiser, 2003a). A leader who values a directive leadership style might demean an approach that encourages decision-making autonomy—or vice-versa. Polarized mindsets can promote lopsided leadership styles (Kaplan & Kaiser, 2003a). *Lopsided* leaders rely on preferred behaviors regardless of their situational suitability; they are vulnerable to overusing strengths to the point of weakness (Kaplan & Kaiser, 2006). Effective leaders learn to use the right behaviors at the right time. Leaders are *versatile* when they “continually adjust their behavior, deftly applying the right approach, to the right degree, for the circumstances at hand. These are people who can pivot readily from forcing a tough issue to fostering harmony” (Kaplan & Kaiser, 2003a, p. 22). The expert leader may intuitively grasp how to balance strengths

without overusing them, but the versatile leader model organizes leadership behaviors such that the dynamic relationship between different approaches can be examined.

Conceptual Structure of the LVI

Traditional instruments measure leadership behaviors as isolated, unrelated aptitudes (Hollenbeck & McCall, 2006). They cannot measure whether behaviors are dynamically balanced (Kaplan & Kaiser, 2003a). Balance is central to the concept of leadership versatility and is reflected through two elements of the LVI design: the Too Little/Too Much (TLTM) response scale and the duality-based construction of the instrument (Kaplan & Kaiser, 2003a; 2006).

Too Little Too Much response scale. On the LVI, 48 behavioral items are measured through the TLTM response scale. Responses to each item in a matched pair are recorded through the Too Little Too Much (TLTM) response format, where values range from -4 to +4, respectively. The scale is bi-directional, incorporating values that range from -4 (*much too little*) to 0 (*the right amount*) to +4 (*much too much*). When leaders use a behavior effectively, raters score them with the central value 0 (*the right amount*). If leaders are deficient in a behavior, then raters score them with negative values ranging between -1 (*a little too little*) and -4 (*much too little*). If leaders overuse a behavior, raters indicate so by scoring them with positive values ranging between +1 (*a little too much*) and +4 (*much too much*). Thus, the TLTM response scale enables raters to simultaneously indicate the frequency of a behavior as well as judge its effectiveness. Therefore, the TLTM response scale may be considered an *evaluation-of-frequency* scale

(Kaiser & Kaplan, 2005). It offers the benefits of both frequency and evaluation scales, but also provides feedback on overused strengths (Kaiser & Kaplan, 2005).

A duality-based structure. Balance is inherent within the duality-based structure of the LVI (Kaplan & Kaiser, 2006). Similar to the Taoist concept of *yin* and *yang*, one side of a duality is the paradoxical complement of its counterpart (Kaplan & Kaiser, 2006). The generic structure of a duality and its components will be described; thereafter a more detailed overview will be provided of each duality and its components.

There are three distinct hierarchical levels within a duality: the duality level, the dimension level, and the sub-dimension level. The duality level describes the focus of measurement. Kaplan and Kaiser identified two dualities, one focused on *how* leaders lead and the other focused on *what* leaders lead (Kaplan & Kaiser, 2006). The Forceful / Enabling duality measures *how* leaders influence constituents through interpersonal processes. The Strategic / Operational duality measures *what* leaders lead—the organizational issues on which they choose to focus (Kaplan & Kaiser, 2006). Each *duality* consists of 24 items.

The 24 items of a duality are evenly divided into two opposing, but complementary *dimensions* (Kaplan & Kaiser, 2006). The dimensional level describes a specific approach to leadership. In the Forceful / Enabling duality, the Forceful dimension is comprised of 12 items that address various aspects of forceful leadership. The Enabling dimension mirrors this structure on the opposite side of the duality.

Dimensions are divided into three *sub-dimensions*, each consisting of 4 items. The sub-dimension level describes specific aspects of leadership; they articulate distinct

behavioral factors comprising each dimension. As with previous levels, every sub-dimension has its complementary opposite, producing a symmetrical structure at the sub-dimension, dimension, and duality levels (see Appendix Z).

The Forceful / Enabling duality. In a literature review that spanned fifty years and nearly thirty dichotomy-based theories, Bass (1990) identified two predominate themes related to interpersonal influence: autocratic use of power (task focus) and democratic use of power (concern for people). In the versatile leader model, the *Forceful dimension* embodied autocratic, task-oriented behaviors and the *Enabling dimension* represented democratic, people oriented behaviors (Kaplan & Kaiser, 2006).

Bass's (1990) survey of literature identified three distinct theoretical perspectives on autocratic and democratic leadership: (1) locus of power, (2) decision-making styles, and (3) motivational orientation (Kaplan & Kaiser, 2006). Using these theoretical perspectives, Kaplan and Kaiser (2006) segmented the Forceful and Enabling dimensions into three sets of sub-dimensions.

Takes Charge versus Empowers Others sub-dimensions. The first pair of sub-dimensions addressed locus of power. *Takes Charge* represented a forceful, self-driven locus of power and *Empowers others* represented an enabling locus of power focused on followers. The LVI used four items to measure *Takes Charge* and four to measure *Empowers others*. Each item had its complementary opposite too. For example, item 1f in *Takes Charge*, “*Takes Charge*—in control of his or her unit” was the complementary opposite of Item 1e in *Empowers others*, “*Empowers* subordinates to run their units—able to let go” (Kaplan & Kaiser, 2006a, p.5).

Declares versus Listens sub-dimension. The second pair of sub-dimensions addressed decision-making. *Declares* represented forceful decision-making processes centered on a leader's independent judgment whereas *Listens* represented a more inclusive and participative decision-making process. As in the previous pair of sub-dimensions, four complementary pairs of items measured *Declares* and *Listens*.

Pushes versus Supports sub-dimension. The third pair of sub-dimensions addressed task versus relationship orientation. *Pushes for performance (Pushes)* represented a forceful orientation with a primary concern for task completion whereas *Supports people (Supports)* represented an enabling orientation with primary concern for people and positive relationships. Four matched item pairs measured *Pushes* and *Supports*.

The Strategic / Operational duality. The Strategic / Operational duality addressed the type of business or organizational matters upon which leaders focus time and attention (Kaplan & Kaiser, 2006). As with the previous duality, the Strategic / Operational duality is composed of 24 items divided into complementary dimensions and sub-dimensions. Kotter (1990) juxtaposed differences between leading and managing. Leaders established visionary direction for their organizations whereas managers planned the execution of these visions. Leaders were change agents who sometimes eschewed structure; managers preserved structure and control. Strategic and Operational dimensions differed along similar lines. The Strategic dimension was similar to Kotter's leadership construct and the Operational dimension was similar to Kotter's management construct (Kaplan & Kaiser, 2006).

The first pair of sub-dimensions addressed timeline and activity. *Long-term direction* represented strategic emphasis whereas *short-term execution* represented an operational emphasis. The second pair of sub-dimensions addressed orientation. *Growth* represented a strategic concern with expanding capability and capacity, whereas *efficiency* represented an operational concern about conserving resources and focusing effort. The third pair of sub-dimensions addressed climate. *Innovation* represented the extent to which leaders encourage creativity and support for employees trying new approaches, whereas *order* represented a more systematic process governed by disciplined procedures (Kaplan & Kaiser, 2006).

Summary of conceptual structure. Responses on the LVI are recorded through the Too Little Too Much (TLTM) response format, where values range from -4 to +4. The most favorable response is at the mid-point, 0 (*the right amount*). Under-rating is scored on the left side of the scale (-1 to -4) and over-rating is scored on the right side of the scale (+1 to +4). The 48 items are symmetrically arranged into two dualities comprised of two, 12-item dimensions. The Forceful / Enabling duality measures interpersonal aspects of leadership, or how leaders influence others. The Strategic / Operational duality measures what leaders tend to focus attention and effort toward. Even though dimensions are conceptually related to one another, they are not polar opposites of a single underlying continuum. Each side of a duality is an empirically distinct construct from its complement (Kaplan & Kaiser, 2006). Dimensions are divided into three pairs of complementary sub-dimensions. Each sub-dimension is linked to its complementary mate through pair-wise relationships at the item level. The empirical

relationship between sub-dimensions and the overall factor structure will be discussed in a psychometric overview of the LVI instrument.

Empirical Support for the LVI Hierarchical Structural Model

The Forceful/Enabling duality, dimensions, and sub-dimensions of the LVI have been empirically supported through a series of structural factor analyses (Kaiser & Craig, 2001; Kaplan & Kaiser, 2006). Kaiser and Craig (2001) explored three possible models for the factorial structure of the forceful/enabling duality: (1) a one-factor model where forceful leadership corresponded to one end of the continuum and enabling leadership on the other; (2) a two-factor model where forceful and enabling leadership were uncorrelated; and (3) a two-factor model where both factors were inversely related, indicating that when one factor is overused, the other will be underused. The two-factor, inversely related model provided the best fit for the data (Kaiser & Craig, 2001; Kaplan & Kaiser, 2006).

The development and evaluation of the structural model occurred in two stages. First, Kaplan and Kaiser (2006) conducted an item analysis and exploratory factor analysis of data collected during a sampling period that ranged from 2001 to 2003 ($n = 941$). They used these analyses to select items that best fit the Forceful/Enabling conceptual structure. Second, using a separate data sample collected from 2003 to 2004 ($n = 2707$), they conducted a confirmatory factor analysis to determine how well the data fit the conceptual model. Results from this confirmatory factor analysis indicated that the seventeen items gleaned from the exploratory factor analyses provided an adequate measurement model (Kaplan & Kaiser, 2006). Items loaded on their intended sub-

dimensions, and these sub-dimensions loaded on their respective forceful or enabling dimension. The forceful and enabling dimensions were inversely correlated ($r = -.58$, corrected for measurement error) (Kaplan & Kaiser, 2006). Therefore, the two-factor, inversely related model provided the best fit for the data (Kaiser & Craig, 2001; Kaplan & Kaiser, 2006). These findings supported the polarity effect—as leaders emphasized one side of a duality, they under-used the other side (Kaplan & Kaiser, 2006).

Though the polarity effect for forceful/enabling dimension was strong and empirical evidence supported a hierarchical structure, the same could not be said for the strategic/operational duality. The correlation between these factors averaged around 0 (Kaplan & Kaiser, 2006). Kaplan and Kaiser (2006) reported low incidence of overdoing Strategic leadership. This restriction of range in scores made it difficult to find evidence for a statistical relationship between strategic and operational dimensions (Kaplan & Kaiser, 2006). Evidence for an inverse relationship has been reported on certain sub-dimensions, the data was not sufficient to support a hierarchical, duality-based model for strategic and operational leadership (Kaplan & Kaiser, 2006).

The strategic and operational duality is not well suited to for leadership development in the majority of student-affairs settings. The constructs address management concepts that exceed students and entry-level professionals' level of experience and competency.

The LVI exhibited strong predictive validity. In a sample of nearly 700 executives, the correlation between high versatility and overall effectiveness was .71 (Kaiser & Kaplan, 2006). That is, versatility accounted for half of the variance in leader

effectiveness. Possessing a “toolkit” of competency-based leadership skills is important—we can assume that executives have a broad array of leadership skills—to be most effective, leaders must learn which strategy to use, when to employ it, and how to avoid using it to excess (Kaplan & Kaiser, 2006).

The Value of the LVI as a Development Tool

As researchers became more adept at identifying and measuring competencies, organizations have used them as key indicators for identifying high potential leaders, developing leaders, and selecting leaders when hiring (Lombardo & Eichinger, 2000). Opponents of competency models claim that they oversimplify the complexity of leadership, that depending on the situation, any strength can become a weakness (Hollenbeck & McCall, 2006). The LVI has a duality-based structure that more accurately captures the tensions and tradeoffs characteristic of leadership. Through the LVI, leaders receive immediate, unambiguous feedback that helps them understand more about their shortcomings, strengths, and strengths overused (Kaiser & Kaplan, 2003; 2006). These insights are especially valuable in cases where leaders must learn to negotiate complex issues and do so with a measure of interpersonal savvy.

Some have argued that leadership development is at an important crossroads for our society (Drucker, 1998; Kaiser, 2005; Tierney, 2006; Weik, 2005). Lombardo and Eichinger (2000) stated that organizations will have greater leadership needs, but fewer personnel capable of addressing those needs. Leadership development is as critical now as it has ever been. Development programs and processes need to evolve by addressing the critical issues leaders face. Leadership is a paradoxical phenomenon that requires

leaders be capable of complex thinking; the right answer is not forceful or enabling, but rather some blend of the two (Martin, 2007). An assessment that reflects this paradox provides greater insight into the nature of these challenges and the stresses they place on leaders (Kaplan & Kaiser, 2006).

To date, the LVI has been used to develop executive-level leaders. The demographic shift brought about as baby boomers retire adds extraordinary complexity to our society's leadership challenges (Drucker, 1998). Leaders need to learn complex leadership skills much earlier in their careers. Unless the LVI can be modified to suit entry- and mid-level leaders, its benefits will be limited to a relatively small population of executive leaders. Such a limitation would be unfortunate. Leadership development is an iterative process (McCauley & Van Velsor (2004). The more one learns the more one is able to learn (Kruger & Dunning, 1999). If students and young professionals can learn more sophisticated techniques earlier in their careers, they have the potential to develop into better leaders, faster.

Developing Future Leaders

Higher Education serves a critical role in developing leaders equipped to address present challenges and those of the future (Astin & Astin, 2000). University mission statements proclaim the importance of developing leaders, providing service to community, and preparing students for the world of tomorrow (Morphew & Hartley, 2006; Smart, Ethington, Riggs, & Thompson, 2002). Astin and Astin (2000) posited that college is both an opportune time and a strategic place to develop leaders. Pascarella and Terenzini (2005) reported that the college experience helps students increase their

leadership abilities through involvement with student government, student organizations, leadership classes, student employment, and other activities where students engage in the activities of leadership. To date, however, student-focused leadership research has been unable to identify specific settings or contexts that contribute to these gains (Pascarella & Terenzini, 2005). Gains are attributed to the cumulative effects of college rather than a specific setting (Pascarella & Terenzini, 2005).

A review of student leadership literature revealed several research methods that could contribute to inconclusive findings. First, researchers have been highly dependent on self-report methodology (Pascarella & Terenzini; Turrentine, 2001). Self report measures of ability and behavioral performance are troublesome on three grounds: First, these types of self-report measures are of dubious validity. Self-report measures of ability and behavior are neither correlated with observer ratings of these characteristics nor objective measures (Beehr, Ivanitskaya, Hansen, Erofeev & Gudanowski, 2001; Conway & Huffcut, 1997). Second, self-report ratings tend to be inflated. Again, compared to observer ratings or objective measures, self-ratings suggest greater ability or more favorable behavioral performance (Harris & Schaubroeck, 1998; Kruger & Dunning, 1992). The third problem with self-ratings and leadership, in particular, is that leadership is a social behavior; followers respond to leaders based on their perceptions of those leaders. A leader's self-perception is of little consequence in this regard (Atwater & Yammarino, 1992).

In a review of student leadership research over the past decade, Pascarella and Terenzini (2005) noted that research on interpersonal behaviors was largely ignored in

favor of examining how institutional characteristics influenced students' self-reported leadership abilities. Researchers quit studying how leaders behaved in favor of studying how institutions impact leaders. Focusing a majority of research on institutional characteristics, while neglecting to balance it with studies on interpersonal behavior, may be likened to studying an ecosystem without considering the relationships among the species within (Antonakis, Schriesheim, Donovan, Gopalkrishna-Pillai, Pellegrini, & Rossome, 2004). Using aggregated data to make inferences about individual-level effects can lead to erroneous conclusions known as ecological fallacies (Antonakis, Schriesheim, et al, 2004). These methodological problems could be limiting the quality of student-affairs leadership research.

Institutional-level research serves a valuable purpose, but its value could be increased if institutional research was complimented by studies of interpersonal behavior. To perform this research, however, multi-rater instrumentation should be employed. Unfortunately, there are few multi-rater instruments designed and validated for the student-affairs context. Kouzes and Posner's (2003) Student Leadership Practices Inventory (SLPI) is the only multi-rater leadership assessment instrument designed for and validated on students (Leadership Challenge, 2008; Posner, 2004). Kouzes and Posner's (2003) SLPI has been used to study the leadership practices of resident advisors (Posner & Brodsky, 1993), fraternity leaders (Posner & Brodsky, 1992; Posner, 2004) sorority leaders (Posner & Brodsky, 1994), and orientation advisors (Posner & Rosenberger, 1998). But the SLPI has limitations related to its frequency response scale that can impact its utility in certain student affairs settings.

Selecting an instrument and matching it with the appropriate context is critically important (Kroeck, Lowe, & Brown, 2004). The information provided by multi-rater instrumentation can be valuable, but certain criteria should be met. Basic knowledge of multi-rater instrumentation fundamentals is important (Leslie & Fleenor, 1998). Multi-rater instruments can be more complicated to administer than self-report instruments because they involve multiple parties. Multi-rater assessments require administrative processes that respect rater confidentiality (Fleenor & Leslie, 1998). These instruments might also provide lengthy feedback reports that can be confusing to the leaders being rated—and potentially confusing to feedback facilitators unfamiliar with debriefing such instruments (Leslie & Fleenor, 1998). They should be used in settings where students and feedback providers have access to good training and supervision.

The housing and residence life setting meets many of the criteria outlined above. Multi-rater assessments are already used to measure a variety of outcomes, including residents' level of satisfaction and residence hall climate (although leadership behaviors are not included in these surveys); therefore residence life administrators and staff members are familiar with this type of instrumentation (ACUHO / EBI, 2008). Furthermore, residence life staff members receive frequent training and ongoing supervision (Winston & Fitch, 1993). Incorporating multi-rater feedback into this process can be a natural extension of current practices.

The assessment context is an important consideration when validating a behavior-based, multi-rater leadership instrument. Ideally, observers and leaders need to have a clear understanding of their respective roles, and have prior experience working

with one another in a formally defined leader-member relationship. The housing and residence-life context provides a number of characteristics that make it attractive for a multi-rater validation study. First, a large number of students occupy formally defined, paraprofessional leadership roles. A *paraprofessional role* is one where students are hired, trained, and supervised to assume responsibilities and perform tasks that promote the development of peers and support a healthy residential living and learning environment (Winston & Fitch, 1993). Both undergraduate students (resident advisors) and graduate students (hall directors) occupy paraprofessional roles (Winston & Fitch, 1993).

Resident advisors, also called resident assistants (RA represents both in this study), are charged with helping to build and maintain an inviting, inclusive residential community. They must also hold student residents accountable to the rules and regulations of the institution (Upcraft & Pilato, 1982). Because of these unique responsibilities, stress levels and potential for burnout is high (Paladino, Murray, Newgent & Gohn, 2005). Therefore RAs require specific training, assistance, and attention (Blimling, 1998; Upcraft & Pilato, 1982).

Hall directors (HDs) serve as direct supervisors for RAs. The HD is an entry-level management position in housing and residence life. Many HD responsibilities are similar to those of a RA, but the scope and depth of responsibility is greater. Whereas RAs manage a floor of residents, for example, the HD is responsible for an entire residence hall. HD responsibilities include management of a residential unit with sizes ranging from 100-500 beds, structuring the educational and psychological environment of the

residence hall, maintaining discipline, facilitating administrative functions, counseling troubled residents, encouraging social and co-curricular programming, and providing training and supervision for RAs (Winston & Fitch, 1993).

Drawing on concepts initially proposed by Lewin and Lippit (1938), Upcraft and Pilato (1982) described the most common leadership styles of paraprofessional residence hall leaders. Some paraprofessionals use an authoritarian style. Authoritarian leaders autocratically set expectations, believe they know the answers, and expect constituents to follow their direction or suffer punishment for disobedience. Authoritarian paraprofessionals have limited success as supervisors (Upcraft and Pilato, 1982). They reported feeling isolated and lonely; their behaviors pushed constituents away. Constituents hide their actions from authoritarian leaders because they feel micromanaged (Winston & Fitch, 1993). Authoritarian leaders emphasize tasks over relationships (Bass, 1990)—on the surface they appear to maintain order, but they do not possess enough relationship equity to be as effective as possible (Upcraft & Pilato, 1982).

A second approach identified by Upcraft and Pilato (1982) was the RA's RA. This type of supervisor is highly relationship focused. These supervisors want to have warm, personal friendships with their supervisees. These supervisors focus on the personal interests of the RA, sometimes excluding the interests of the institution or student residents. If an RA had a personal problem that continually impinged on his ability to serve residents, the "RA's RA" might try to attend to the RA's personal needs but would have difficulty holding the RA accountable for the institutional responsibilities incumbent with their role.

Some paraprofessionals utilize an uninvolved approach to leadership. This style has been identified as laissez-faire supervision (Upcraft & Pilato, 1982). Laissez-faire leaders remain hands-off until an intervention is absolutely necessary. The leader does not try to build relationships or provide structured guidance into task management. RAs often feel under prepared and under trained working for this style of leader (Upcraft & Pilato, 1982).

Winston, et al., (1984) identified a fourth type of supervisor: the *synergistic supervisor*. Synergistic supervisors are adept at managing the paradoxes of supervision. They respect the needs of individual supervisees, but these needs are considered in the context of the institutional mission. The synergistic supervisor negotiates a balance between the task-focused, authoritarian style and the relationship focus of the RA's RA style. This balance enables the supervisor to maintain a cordial relationship with supervisees, yet hold them accountable when necessary. Winston, et al., (1984) believed the synergistic supervisor was the most effective approach to supervision. Winston and colleagues designed a supervision model, the synergistic supervision model, to help paraprofessional and professional student affairs supervisors learn to utilize a more balanced, or "synergistic" leadership style (Winston, Ullum, & Werring, 1984; Winston & Creamer, 1997; 1998; Winston & Hirt, 2003).

The synergistic supervision model has expanded in complexity and is currently used to supervise all levels of student affairs professionals, not merely entry-level supervisors (Janosik, Creamer, Hirt, Winston, Saunders, & Cooper, 2003). In a study of high quality supervisors, Armenio and Winston (2001) cataloged the characteristics of

effective supervision. Synergistic supervisors provide structure and direction, but also manage to encourage input from subordinates. They make their points clearly and unobtrusively, but listen to supervisees' perspectives. Supervisors communicate high levels of expectation, but convey a sense of support (Armenio & Winston, 2001). When supervisors were asked to describe how they managed to strike this synergistic balance, many confessed they did not know or suggested they learned through trial and error (Armenio & Winston, 2001). Some specifically noted that they did not learn this skill in graduate school. Many said they learned synergy by not doing the things that frustrated them when they were supervisees. Given the importance of effective supervision, Armino and Winston (2001) claimed that more and better preparation was needed to help supervisors to learn how to develop synergistic relationships.

The synergistic supervision model articulates a sophisticated, integrated perspective of leadership (Winston & Creamer, 1997; Winston & Fitch, 1993; Winston, Ullom, & Werring, 1984), but it does not provide a practical method for quantitatively measuring "synergy." Creamer and Janosik (2003) recommended using traditional frequency and evaluation scales to measure the behavioral performance of synergistic supervisors, but such instruments are unable to measure balance because they cannot measure overuse of a behavior (Kaplan & Kaiser, 2006). The SLPI is similarly limited in this respect. Kaplan and Kaiser's (2006) versatile leader model and the LVI provide a solution for this shortcoming. Both the versatile leader model and the synergistic supervision model acknowledge the paradoxical nature of leadership. The synergistic

model provides a broad conceptual overview and the LVI quantitatively measures the underlying behaviors necessary for putting synergy into practice.

To date, no quantitative, multi-rater studies on a housing and residence life population have researched versatility or the synergistic supervision model. A limited number of studies have investigated HD leadership and how their leadership style impacts RA outcomes. Komives (1991b) examined the relationship between HD interpersonal leadership styles (transformational versus transactional), RA levels of satisfaction and motivation, and RA perceptions of their HD's effectiveness as a leader. The most effective HDs emphasized an engaged, transformational, leadership style characterized by high consideration for others, charisma, and intellectual stimulation (Komives, 1991b). Given the stressful nature of resident advising and the priority resident advisors place on supportive relationships, the preference for more engaged, transformational leadership styles was not surprising (Komives, 1991b). A self versus observer ratings comparison indicated potential blind spots for HDs. In comparison to RA perceptions, HDs overrated themselves on the amount of positive, highly engaged leadership behaviors they provided. HDs underrated themselves on less engaged leadership characteristics (e.g., remaining distant unless a problem occurred) or non-leadership behaviors (Komives, 1991b). HDs also overestimated RA satisfaction and motivation levels in comparison to levels reported by RAs (Komives, 1991b). This disparity between self-report and other-report data is not unique; in fact, studies within business and industry have demonstrated similar disparity (Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski, 2001; Conway & Huffcut, 1997; Harris & Schaubroeck, 1998,

2000). It seems possible, then, that the manner in which leadership data is collected (i.e., self-report versus other-report) is critical.

Researchers have considered how factors such as gender might influence the leadership process between HDs and RAs. For example, Komives (1991c) explored the degree to which gender pairings of RAs and HDs affected RAs level of satisfaction with HD leadership, overall job satisfaction, and motivation. Gender pairings were not found to significantly impact leadership outcomes in any of the four combinations studied (Komives, 1991c). Komives (1991c) noted these findings supported previous studies by Bartol and Wortman (1976) and Stitt, Schmidt, Price, and Kipness (1983).

Komives (1991a) examined the relationship between HDs self and other (RA) reported leadership factors and achievement styles (also self and other-reported). Male HDs who rated themselves high on transformational characteristics attributed achievement to their ability to take charge and direct or control others. Female HDs held a significantly different perspective on transformational leadership and the characteristics that contributed to their transformational style. Females thought their relational achievement style (vicarious, contributory, and collaborative) contributed to their transformational qualities. When HDs leadership and achievement styles were rated by RAs, however, a common pattern emerged for both males and females: all HDs who were perceived to be transformational used collaborative, relational leadership styles. This perception contrasted sharply with Male HD self-reported achievement and leadership styles (Komives, 1991a).

Since RA satisfaction levels (Komives, 1991b), stress levels, and burnout levels can be related to HD leadership behaviors (Paladino, Murray, Newgent, & Gohn, 2005), increasing our ability to understand the leader-member relationship between HDs and RAs can provide multi-faceted benefits for departments of housing and residence life. HD training and development may be enhanced through this knowledge. A training program that incorporates systematic feedback processes can help HDs learn how their leadership is perceived, recognize their leadership tendencies, and the impact of those tendencies. If this program incorporated principles of synergistic supervision (Winston, et al., 1984), then the versatile leader model could help them better understand how to manage between forceful leadership and enabling leadership, and thus be a more effective synergistic supervisor.

These findings emphasize the vital need for leadership assessment and feedback in a campus-based setting. Unfortunately, however, the vast majority of leadership assessment instruments have been developed and validated on professional managers (e.g., Bass & Avolio, 1995/2000; Kaplan & Kaiser, 2006) and there are few leadership instruments that have been designed for college campus applications. The few leadership instruments that have been developed for college populations have design limitations that hinder their ability to measure synergistic supervision. An LVI developed for a campus-based context can help alleviate this problem.

Statement of the Problem

Leadership is not the product of an individual, but rather the result of collaborative group processes. That is, leadership is a social phenomenon. When leader-

member relationships are an emphasis of study, relationships should be measured with instruments capable of assessing them from multiple perspectives, not merely from self-report sourced data (Antonakis, Cianciolo, & Sternberg, 2004). Furthermore, self-reported data are not correlated with objective measures nor with observer reports (Beehr, Ivanitskaya, Hansen, Erofeev & Gudanowski, 2001; Conway & Huffcut, 1997). Observer reports of leadership behavior have demonstrated significantly greater validity and reliability than self-reported data (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988). Therefore student-focused research should include multi-rater methods that are focused on behavioral observation.

To better understand HDs leadership behavior and the dynamic elements of HD/RA relationships, then multi-rater instruments should be incorporated into the assessment process (Winston & Fitch, 1993). Kouzes and Posner's (2003) Leadership Practices Inventory Student Version (SLPI) is a multi-rater assessment instrument. By virtue of its multi-rater design, the SLPI is better suited to studying relational leadership. But the SLPI also has weaknesses. The response scale requires raters to indicate frequency of behaviors observed on a Likert-type scale that ranges from one (*rarely observed*) to five (*frequently observed*) (Posner, 2005). Although frequency scales are the most popular type of scale used in multi-rater assessment (Leslie & Fleenor, 1998), Shipper (1991) reported that increased behavioral frequency may exist independent of behavioral mastery. The SLPI cannot indicate whether a leader relies too much on a narrow range of behaviors.

Because the SLPI is unable to measure strengths overused, it is limited in its ability to provide feedback that can help HDs become more versatile, synergistic leaders. A student-focused version of the Leadership Versatility Index would be a better match for the skills and aptitudes required of synergistic supervisors. To date, no such instrument exists.

Purpose of the Study

The overarching purpose of the current study is the construction and validation of a new, multi-rater, student leadership assessment instrument named the Leadership Versatility Index-Student (LVI-S). The construction phase of this study involves several steps. First, a literature review explores elements of the versatile leader model and demonstrates how leadership versatility is relevant to synergistic supervision and the leadership development leaders in housing and residence-life settings. Step two of the construction phase outlines procedures taken to modify the executive-focused LVI (Kaplan & Kaiser, 2006) for use with college students. The authors of the LVI granted permission to modify this instrument for a college population and a copy of this written permission is included in Appendix H. The result of these modifications will produce the initial version of a student focused, multi-rater leadership instrument named the Leadership Versatility Index-Student (LVI-S) that will be administered to a sample of Hall Directors (focal leaders) and Resident Assistants (subordinates). Multi-rater instruments consist of two separate, but parallel survey forms. The LVI-S “instrument” refers to both the self-rating form and the observer rating form. The “self” version of the survey asks the focal leader, in this case the HD, to rate her or himself. The “observer”

survey form asks a parallel set of questions where subordinates (RAs) are asked to rate their perception of the HD's leadership behavior.

Research Questions

This study is designed to answer the following research questions:

1. Does the LVI-S verify that there are two leadership factors: forceful leadership and enabling leadership?
2. What proportion of the HD reputational effectiveness can be explained by the six LVI-S sub-dimensions? Are any of the six sub-dimensions statistically significant predictors of HD reputational effectiveness as reported by RAs?
3. To what extent do scores on the LVI-S and the SLPI correlate, providing evidence of convergent construct validity of the LVI-S?
4. Do most HDs underestimate or overestimate their F/E dimension scores in comparison to RAs ratings of the HD's F/E dimension scores?
5. Is there a statistically significant mean difference in effectiveness ratings between HDs who overestimate their F/E dimension scores versus HDs who underestimate their F/E dimension scores?
6. Which regression model better predicts reputational effectiveness scores: Model 1, an additive model where interaction between complementary F/E sub-dimensions is not accounted for, or Model 2, a model that incorporates the focal leader's joint standing scores (versatility) on complementary F/E sub-dimensions?

Need for the Study

The leadership challenges of society are increasing daily. Leadership problems are growing more complex (Friedman, 2007), and the leaders that will be tackling these problems will have less experience than their predecessors (Teegarden, 2004; Tierney, 2006). Leaders must develop an increased capacity to manage paradox (Martin, 2007). They need to recognize that effective leadership is focused on both the individual and the institution (Winston, Ullom, & Werring, 1984), and learn how to pivot between forceful and enabling leadership behaviors (Kaplan & Kaiser, 2006).

The development and validation of the LVI-S will provide housing and residence life administrators with a tool that measures versatility, a key component of the synergistic supervision model. Without the LVI-S, there are no student-focused measurements sensitive to overuse of strengths, nor any that can measure the dynamic balance of strengths. Absent a measure of excess, there is no way to determine if leaders are learning to use leadership skills as effectively as possible (Kaplan & Kaiser, 2006).

The LVI-S will provide residence hall administrators the ability to measure how HDs balance the inevitable tensions and tradeoffs characteristic of residence hall management. Designing an instrument that measures HDs' forceful and enabling leadership behaviors can provide valuable insights that may inform the development of HD leadership training and assessment programs. Effective training programs can help HDs learn to be more synergistic supervisors and enable them to provide higher quality supervision (Upcraft & Pilato, 1982; Winston & Fitch, 1993). Effective training has been identified as a factor to minimize RA burnout (Paladino, et al., 2005). Furthermore, an

instrument such as the LVI-S can help housing and residence life administrators establish pre and post benchmarking metrics to demonstrate whether training programs are producing their intended effects.

Key Terms and Definitions

Enabling leadership—Creating conditions for others to take the lead; empowering others, delegating, supporting, making it easy for others to push back and express their opinions on how to accomplish tasks (Kaplan & Kaiser, 2006)

Forceful leadership—leadership style where a leader actively directs a course of action; a leader actively serving as a force of influence (Kaplan & Kaiser, 2006)

Leadership behaviors—observable actions of leadership (Sashkin, 2004).

Leadership Practices Inventory-student version (SLPI)—Student-focused multi-rater instrument used to measure the Five Practices of Exemplary Leadership (Posner, 2004).

Lopsided leadership—the opposite of versatile leadership. Lopsided leaders are biased toward using skills on one side of a leadership duality versus the other. This imbalance of skills leads toward overuse of a strength to the point of weakness (Kaplan, 1996).

Multi-rater assessment—Assessment method wherein multiple raters, called observers, rate a leader. This method is often termed *360-degree assessment* because different types of observers can provide unique perspectives of the subject which, when combined, can provide a well-rounded course of feedback (e.g., a supervisor provides a supervisory perspective, a peer provides a peer perspective, etc.) (Leslie & Fleenor, 1998).

Self-report assessment method—assessment method where individuals are asked to rate themselves on specific measurement criteria. Self-report methods are most helpful when used to measure variables that are verifiable through multiple sources of data (e.g., demographic information) (Podsakoff & Organ, 1996).

Synergistic supervision— the synergistic supervisor models how to balance between concern for the individual and maintaining respect for institutional objectives; this style of supervision requires supervisors to manage paradox effectively and effectively balance forceful and enabling behaviors (Winston, Ullom, & Werring, 1984).

Versatile leadership—a balanced approach to leadership wherein the leader possesses a wide repertoire of leadership strengths, knows how and when to appropriately use these skills, and does not overuse strengths to the point of detriment (Kaplan & Kaiser, 2006).

Brief Overview

This study is presented in five chapters. The first chapter has provided an introduction to leadership assessment, leadership versatility, and the rationale for the LVI-S. The purpose of the study, statement of the problem, and need for the study are outlined in this introduction. In addition, definitions of key terms are provided. The second chapter contains a review of the literature as it relates to leader development in residential college campus settings, leadership assessment, the synergistic supervision model, and leadership versatility. The third chapter includes the methodology that will be used in the study, including participants, sampling method, instruments, and data analyses. A description of multi-rater administration procedures, the scoring process, and development and preliminary validation to date are presented. The fourth chapter presents

the results of this research according to each research question. Finally, the fifth chapter will summarize the study and include limitations and recommendations for future research in the area of leadership versatility in collegiate settings.

CHAPTER II

LITERATURE REVIEW

The Contemporary Leadership Development Context

Great leadership benefits the collective, whether at the level of the country, a community, or an organization (Hogan & Kaiser, 2005). Conversely, the consequences of poor leadership can be devastating. Hogan and Kaiser (2005) cited the example of Foday Sankoh, former dictator of Sierra Leone. Sankoh was charming and charismatic—and able to mobilize his following. During the 1980s and 1990s he savagely pillaged his country. Rabid followers destroyed families and looted the country's resources; by the turn of the century Sierra Leone was acknowledged as the poorest country in the world (Hogan & Kaiser, 2005). Leadership is one of the most highly scrutinized processes in the social sciences (Antonakis, Cianciolo, & Sternberg, 2004). Leadership matters—and warrants the close scrutiny it receives (Dionne, Yammarino, Atwater, & James, 2002).

Astin and Astin (2002) argued that the quality of leadership in the United States is in decline. Recent history *Supports* this observation from financial, civic, and demographic perspectives. In the last decade, the Federal deficit has doubled from \$5 trillion to more than \$10 trillion (U. S. Treasury Department, 2008); the gap between the wealthiest and poorest Americans continues to grow (University of Michigan, 2007); Congress granted Wall Street a \$700 billion economic bailout because of poorly regulated loan practices (Krauthammer, 2008; New York Times, 2008). Combined with

limited regulation from Washington, these loan practices have contributed to a crisis of confidence in financial institutions. Lines of credit have been severely restricted, stalling the economy (Kodres, 2008). Wall Street mismanagement has impacted Main Street America. Mortgage foreclosures reached record levels (Hilzenrath & Elboghdady, 2007) and November 2008 reports indicated 533,000 new applications for unemployment (Bureau of Labor Statistics, 2008).

Kaiser, Hogan, and Craig (2008) argued that leadership effectiveness should be defined and evaluated by the performance of the group or team for which the leader is responsible. If the financial markets and the economy are any indication, it appears Astin and Astin's (2002) claims are well founded. These leadership problems extend into civic issues as well. Regional and national studies of volunteerism and civic engagement reported decreasing levels of involvement, lower levels of trust in government, and increased apathy about an individual's ability to create positive change (Putnam, 2000). The National Conference on Citizenship (NCC) (2006) indicated these civic engagement trends have persisted for nearly three decades. These results corroborated findings from earlier reports as well (NCC, 2006).

Demographic concerns also are an issue. As baby-boomers retire, communities and organizations may be confronted with a leadership vacuum (Annie E. Casey Foundation, 2004; Lombardo & Eichinger, 2000; Tierney, 2006). This vacuum can be observed on multiple fronts. For example, in a study sampling 115 Human Resource executives representing U.S.-based medium and large corporations, Weik (2005) reported that 50% expect to lose half of their senior managers by the year 2010, 15% anticipate

losing at least three-fourths or more of their senior managers, and 75% reported they were only “somewhat confident” or “not at all confident” about their organization’s internal ability to develop enough leaders to meet future demands. Reports such as these have led some analysts to warn that America is approaching a leadership crisis. Kaiser (2005) noted:

Most organization charts and succession maps are noteworthy for the number of blank slots five years out. And the graying and impending retirement of baby boomers coupled with steep and steady declines in skilled entrants to the workforce adds up to an even greater shortage in the U.S. labor market of the early twenty-first century. This could make the “War for Talent” of the late 1990s look more like a street fight. (p.1)

This leadership void is not restricted to corporate business. Teegarden (2004) conducted a study of 2200 executive directors in the nonprofit sector and found similar results. The Baby Boom generation comprised nearly 75% of all nonprofit leaders. These leaders are expected to transition out of the sector in two waves—the first is ongoing and should continue through 2010, the second should peak in 2020 as Baby Boomers approach traditional retirement age (Teegarden, 2004). If current trends continue over the next decade, nonprofit organizations will need to attract over 640,000 new senior managers—2.4 times the number currently employed (Tierney, 2006). By 2016, nonprofit organizations will need almost 80,000 new senior managers per year (Tierney, 2006). To put this demand into context, Tierney observed, “attracting the required number of managers will be equivalent to recruiting over 50% of every MBA graduating class, at every college and university across the country, every year for the next ten years”

(Tierney, 2006, p. 6). This impending deficit underscores an urgent need for more leaders, and prompts the question “where will these leaders come from?”

Are Leaders Born or Made?

Progression of leadership theories. The practice of leadership dates back to antiquity (Bass, 1990), but the systematic, empirical study of leadership began in the early 20th century (House & Aditya, 1997). Research focused on leaders who were considered “Great Men” and were born with “the right stuff” (Antonakis, Cianciolo, & Sternberg, 2004, p. 33). Researchers explored dispositional characteristics, or traits, that distinguished leaders from non-leaders. Trait theory posited that leaders possessed a unique set of inborn characteristics enabling them to command. Studies from Mann (1959), Stogdill (1948), Gibb (1947) and Jenkins (1947) found traits related to leader effectiveness—but the state of leadership science prevented effective replication of these studies (House & Aditya, 1997). Consequently, trait theory fell out of favor by many leadership scholars (Antonakis, Cianciolo, & Sternberg, 2004; House & Aditya, 1997).

Style and behavioral theories. The trait movement was overtaken by a behavioral focus on leadership style. House and Aditya (1997) cited three major schools, or emphases, of behavioral study. Bales and associates at Harvard University (Bales, 1954) conducted laboratory experiments and observed leader behaviors. Stogdill and Coons (1957), conducted factor analytic studies at Ohio State; and finally, the University of Michigan conducted interview-based studies of behavior (Kahn & Katz, 1953; Likert, 1961; Mann, 1965). Each of these schools converged on two classes of leadership

behaviors: task oriented behaviors and person-oriented behaviors (Bass, 1990; House & Aditya, 1997).

A voluminous body of behavior-focused research followed the emergence of task-oriented and relationship-oriented constructs (Bass, 1990). Researchers sought to identify linear behavioral patterns that were significantly related to performance outcomes, but studies failed to yield universal patterns. Blake and Mouton's managerial grid, for example, hypothesized that the most effective leaders would score high on both task and relationship behaviors. Northouse (2004) noted several studies that supported this proposition (Blake & McCauley, 1991; Misumi, 1985). But Yukl (1994) reported only limited support for a universal high-high leadership style (Northouse, 2004). In fact, the evidence was largely contradictory and inconclusive (Bass, 1990; House & Aditya, 1997; Yukl, 1994). Other variables appeared to moderate leader behavior and outcomes (House, 1971; Fiedler, 1967).

Contingency theories. Researchers speculated the relationship between behaviors and effectiveness was contingent upon situational variables (Antonakis, Cianciolo, & Sternberg, 2004). This finding led to the development of contingency leadership theories. Contingency theories indicated that leader effectiveness is a product of behaviors as well as situational factors (House & Aditya, 1997). Successful leadership outcomes depended on how well the leader's style fits situational needs of the setting and of the people involved (Northouse, 2004). Support for these theories was mixed, but they did advance the understanding of leadership.

Contingency theories focused on the interaction between a leader's personal characteristics and situational variables (Fiedler, 1968), a leader's decision-making processes (Vroom & Yetton, 1973), and a leader's ability to assess a situation and help followers envision a path toward accomplishing their goal (House, 1971; House & Mitchell, 1974). But for every answer that emerged, multiple questions were generated (Antonaikis, Cianciolo, & Sternberg, 2004). Several meta-analyses validated Fiedler's contingency theory (Strube & Garcia, 1981; Peters Harke & Poleman, 1985), but this support was debated (Fiedler, 1995; House & Aditya, 1997). Leader personality and situational factors influenced outcomes, but the theory was limited in its ability to articulate specifically why task motivated leaders succeed in some situations and relationship motivated leaders succeed in others (Fiedler, 1993; Fiedler, 1995).

Complications affected other contingency theories as well. Vroom and Yetton (1973) believed managers' decision-making processes were essential to leadership effectiveness. Research indicated that managers unwittingly limited the criteria they considered when making decisions. Vroom and Yetton (1973) developed a model to expand managers' decision-making processes. The model provided value as a research exercise, but it was cumbersome, complex, and not widely applied (Bass, 1990). In path-goal theory, House (1971) approached contingency theory from the perspective of subordinate motivation and performance. Leaders were responsible for assessing a situation and illuminating paths that would lead followers toward their goals. Path-goal theory was highly complex and difficult to fully test or adequately apply in an organizational environment (Northouse, 2004). Validation studies provided partial

support (House & Mitchell, 1974; Schriesheim & Schriesheim, 1980; Woffard & Liska, 1993), but path-goal theory incorporated so many elements that research has yet to sufficiently address each one (Northouse, 2004). Contingency theory demonstrated that leaders needed to be mindful of how they fit within situations (Fiedler, 1968), that their decision-making processes were significantly related to effectiveness (Vroom & Yetton, 1973), and that follower motivation was contingent on how well a leader could read the situation and the needs of constituents (House, 1971). On the other hand, contingency models were hampered by excessive complexity and by findings that demonstrated validity, but lacked specificity.

New schools of study. Excessive complexity contributed to the decline of contingency research in the latter part of the 1980's (Antonakis, et al., 2004; House & Aditya, 1997), but it helped spawn several new schools of study: relational leadership, information-processing, and transformational theory (also known as neo-charismatic theory) (Antonakis, et al., 2004). Each of these schools delves deeply into specific elements of contingency theory.

Relational theories, such as leader member exchange theory (LMX theory) (Graen & Uhl-Bien, 1995), studied the quality of relationships between leaders and followers. Information processing theory examined how leader cognitions related to behavior (Antonakis, et al, 2004), how followers attributed legitimacy to a leader (Lord, Foti, & DeVader, 1984) and the different ways leaders and followers assigned credit or blame for their performance (DeVader, Bateson, & Lord, 1986). Transformational theory became one of the most popular and influential theories of the last two decades (Lowe & Gardner,

2000). Transformational leadership theory examined how leaders motivated followers. Rather than simply studying transactional exchanges between leaders and followers (do X to get Y), transformational leadership explored how leaders captured followers' hearts and minds and appealed to their moral sensibilities (House, 1977; Burns, 1978; Bass, 1985a).

Research advancements. The emergence of more refined theories altered the research landscape. Transformational theory, in particular, marked a transition from empirical skepticism toward one of reinvigorated excitement (Antonakis, et al., 2004). According to Lowe and Gardner (2000) nearly one-third of the 88 articles published in *Leadership Quarterly* between 1990 and 1999 were based on transformational leadership. Bernard Bass' (1985b) Multifactor Leadership Questionnaire (MLQ) operationalized transformational constructs and served as a cornerstone for transformational research and development (Lowe, Kroeck, & Sivasubramaniam, 1996). The MLQ has been the most prolific transformational instrument, but not the only one. Kouzes and Posner (1987) created the Leadership Practices Inventory, and Sashkin (1984) created the Leader Behavior Questionnaire (LBQ) (Sashkin, 2004) and Podsokoff, MacKenzie, Moorman, & Fetter (1990) created the Transformational Leader Behaviors Inventory (TBI).

Advances in methods, analyses, and instrumentation. Kroeck, Lowe, and Brown (2004) catalogued difficulties associated with the measurement of leadership. Leadership is amorphous, lacking a specific definition that cuts across theories, disciplines, and practices (Kroeck, et al, 2004). The tradition of leadership measurement has not provided a formal analysis of assessment methods nor directly compared assessment methods.

Measurement processes and the rationale underlying these processes have been vaguely defined (Kroeck, et al, 2004). As one might expect, this haphazard approach to methodology has a detrimental impact on the validity of results. Validity was compromised by a host of issues, including improper aggregation of data, inappropriate assessment devices, and the use of measurement scales with poor psychometric performance.

The late 1980s and 1990s were periods of rapid evolution. Advances in theory were complemented by advanced research methods and analytical procedures (Lowe & Gardner, 2000). Examples of methodological advancements include the following: Mabe and West (1982) articulated conditions necessary to maximize validity of self-reported data. Podsakoff and Organ (1986) addressed problems and prospects of self-reported research in organizational data collection. Levels of analysis proved to be another critical factor to consider when selecting a method and process of analysis (Yammarino, Dionne, Chu, and Danserau, 2005). As these and other issues were identified and controlled, threats to validity were managed more effectively (Kroeck, et al., 2004; Podsakoff & Organ, 1986).

Leadership measurement instruments improved as well. Leslie and Fleenor (1998) reported that in 1978, only 24% of the instruments reviewed by the Center for Creative Leadership had received psychometric scrutiny. In 1991, the percentage grew to 40%, and by 1998, 53% met minimum test development standards specified by the American Psychological Association's *Standards for Educational and Psychological Testing* (APA, 1985). Multi-rater instruments increased in popularity (Leslie & Fleenor, 1998). Multi-

rater data were more reliable and valid than self-reported data (Harris & Schaubroeck, 1988; Conway & Huffcutt, 1997; Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski, 2001).

As the science of leadership improved, so did the utility of its findings. In contemporary research, trait theory (previously discarded because of methodological limitations), information processing theory, transformational theory, and newer iterations of contingency theory (referred to as contextual theories) are among the most active research agendas (Antonaikis, et al., 2004). Each one has contributed significant findings toward contemporary literature (Antonakis, et al., 2004). These findings have greatly influenced the leadership development process.

Developing Leadership in Corporate Environments

Competency Models

Improvements in instrumentation, methods, and analysis led to more reliable and valid research outcomes. As social scientists improved research practices, they were able to more accurately identify and measure the characteristics of effective leader (Kroeck, Lowe, & Brown, 2004). Such characteristics were called competencies, defined as measurable characteristics related to work success (Boyzatis, 1982; Lombardo & Eichinger, 2000). Competencies were arranged into competency models, which organizations used for designing training programs, development initiatives, and succession planning (Lombardo & Eichinger, 2000). Competencies included skills, attitudes, and personal attributes such as intelligence (Lombardo & Eichinger, 2000). Organizations identified competencies believed to be required for success in a role.

Competency-based training curricula serves to enhance employee performance in current roles as well as prepare them for future advancement; competency-based assessment is used to evaluate current performance, and competencies help inform employee selection procedures (Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski, 2001; Lombardo & Eichinger, 2000).

Derailment behaviors and metacompetencies. Some competency models include derailment behaviors (Lombardo & Eichinger, 2000; McCall & Lombardo, 1983). Whereas competencies represent the skills or characteristics related to successful leadership, derailment behaviors can impede successful leadership, causing career stagnation, demotion, or firing (Lombardo & Eichinger, 2000; McCall & Lombardo, 1983). As leaders move into different roles at different levels of responsibility within an organization “strengths become weaknesses” (McCall & Lombardo, 1983, p. 11). Sometimes a skill or ability that worked well in one context becomes problematic in another (Charan, Drotter, & Noel, 2002; Lombardo & Eichinger, 2000; McCall & Lombardo, 1983). Lombardo and Eichinger (2000) stated that most derailment behaviors result from inappropriately used strengths. Therefore it is not sufficient to merely develop a competency, one must also know how to use it effectively.

A metacompetency is a skill, characteristic, or aptitude—such as the ability to read—that affects proficiency in other competencies (Briscoe & Hall, 1999; Hall & Moss, 1998). Briscoe and Hall (1999) observed that adaptability, or the ability to adapt one’s skills to the needs of changing circumstances, could be considered a metacompetency. If every competence can also be an incompetence (Hollenbeck &

McCall, 2006), then knowing how and when to apply each ability is critical. This topic has been investigated under the headings of behavioral flexibility (Zaccaro, Foti, & Kenney, 1991; Zaccaro, Gilbert, Thor, & Mumford, 1991), behavioral complexity (Quinn, 1988; Denison, Hooijberg, & Quinn, 1995), integrative thinking (Martin, 2007), leadership agility (Lombardo & Eichinger, 2000), adaptability (Briscoe & Hall, 1999), and leadership versatility (Kaiser & Craig, 2001; Kaiser, Lindberg, & Craig, 2007; Kaplan, 1996; Kaplan & Kaiser, 2003a; Kaplan & Kaiser, 2003b; Kaplan & Kaiser, 2006).

Problems with competency models. Competency models may be helpful, but they can oversimplify the complexity of leadership (Hollenbeck & McCall, 2006). Effective leadership requires constant assessment of contextual, situational, and interpersonal variables; leaders must then act or react appropriately. No list of characteristics and behaviors, regardless of how extensive, can truly represent leadership in action (Hollenbeck & McCall, 2006). Competencies that work well in one scenario might be inappropriate for the next. Effective leaders adjust their style to accommodate these dynamic leadership challenges, but assessment instruments are hard-pressed to account for such adaptability or versatility (Hollenbeck & McCall, 2006).

Measuring Competency Models: Multi-rater Response Formats

Most competency-based, developmentally focused multi-rater instruments use a continuous response scale, or format, similar to the one Rensis Likert (1932) used to measure attitudes (Kaiser & Kaplan, 2005). A continuous scale means that each response has a number assigned, such as 5 = *Always*, 3 = *Sometimes*, and 1 = *Never*, and scores are

calculated by adding or averaging across items in a scale (Leslie & Fleenor, 1998). The 5-point version remains the most popular, but some range as high as 10 points (Leslie & Fleenor, 1998). In leadership applications, these response scales are used to measure behavioral performance. Behavioral variables are used because discrete behaviors are (1) readily observable and measurable and (2) because research demonstrated that behaviorally focused feedback is more likely to result in performance improvements (Kluger & DeNisi, 1996).

The frequency response format. In most cases these response scales take one of two forms: frequency scales or evaluative scales (Kaiser & Kaplan, 2005; Leslie & Fleenor, 1998). *Frequency scales* prompt raters to consider, from “less to more,” how frequently they recall a leader exhibiting specific behaviors (e.g., *never, a little bit, sometimes, often, always*). A second type of frequency scale permits raters to indicate how accurately behavioral statements represent a leader (e.g., *not at all, somewhat characteristic, highly characteristic*) (Kaiser & Kaplan, 2005). Frequency scales excel at detecting leadership shortcomings (Kaiser & Kaplan, 2005). When leaders rarely utilize a skill, raters observe the behavior less frequently and should report lower scores; these lower scores draw attention to behavioral shortcomings.

The evaluation response format. Evaluative scales, on the other hand, ask respondents to impart judgment on a leader’s behavior. Evaluative response scales may be absolute or relative (Kaiser & Kaplan, 2005). Absolute scales prompt raters to consider, in absolute terms, how well a leader performed (e.g., *poor, somewhat effective, outstanding*) (Kaiser & Kaplan, 2005). Relative scales ask raters to consider a leader’s

performance relative to others (*among the worst, average, among the best*) (Kaiser & Kaplan, 2005). Evaluative scales illuminate leadership strengths (Kaiser & Kaplan, 2005). High ratings indicate a leader's behaviors are perceived to be effective with regard to ability in question (Kaiser & Kaplan, 2005). In summary, shortcomings and strengths represent key focus areas for developmental research and feedback, but there is a third focus area these scales are unable to address.

Blind spots in measurement formats. Frequency scales and evaluative scales complement one another; one reports on deficiencies and the other reports on strengths. But neither scale distinguishes between appropriate use and excessive use of behaviors. These scales are blind to overuse of strengths (Kaiser & Kaplan, 2005).

Blind spots in the frequency format. Frequency scales are confounded at the higher ranges. As ratings increase, observers are unable to discriminate between appropriate and overly frequent use of a behavior (Kaiser & Kaplan, 2005). On the 1-5 response format (ranging from *never*, to *sometimes*, to *often*) raters cannot distinguish between the behavior that occurs often and that which occurs *too* often (Kaiser & Kaplan, 2005). Sometimes less is more, but the 1-to-5 response format can inhibit this feedback. Kaiser and Kaplan's (2005) contentions have been stated elsewhere. Yukl (1989) observed that most questionnaire-based leadership instruments measured quantity (frequency) of behaviors but did not address quality of behaviors. Yukl (1989) recommended that future research should address this discrepancy. Shipper's (1991) study of professional managers revealed leaders of high performing units exhibited lower

overall behavioral frequency than the managers of lower performing units (Shipper, 1991). Therefore, an increase in behavioral frequency did not imply behavioral mastery.

Blind spots in the evaluative format. Evaluative ratings become unclear at the low end of the scale (Kaiser & Kaplan, 2005). When evaluative ratings are high, the recipient knows to continue doing what she has been doing. But if a leader is rated “adequate” or “poor” on certain leadership competencies, there is little to indicate why the low rating was given. The score indicates improvement is needed, but it does not provide specific feedback on what to improve. Is the leader deficient in the ability or using it too much? When feedback recipients lack clarity about why lower scores were given, they have greater difficulty correcting their behaviors (Kaiser & Kaplan, 2005).

Implications of blind spots. Measuring leadership competencies through frequency and evaluative scales should not be mistaken with measuring competent leadership (Hollenbeck & McCall, 2006). Competent leadership involves balancing leadership skills with the needs of a situation (Hollenbeck & McCall, 2006). Frequency scales isolate deficient behaviors, but they are unable to distinguish between the right amount of a behavior and too much of a behavior. They are blind to overused strengths. Feedback recipients are at risk for interpreting “high” scores as indicators of effective leadership. Leaders may demonstrate the acquisition of a skill, but that does not mean they know how to apply it effectively. Frequency does not indicate mastery (Shipper, 1991). Considering most derailment variables are overused strengths, this limitation is significant (McCall & Lombardo, 1983; Lombardo & Eichinger, 2000). Evaluation scales provide insight into effective use of leadership strengths but provide little of value when

feedback indicates poor performance. When ratings are fair or poor is it because leaders were deficient in ability or because they overused their strengths? When attributional ambiguity is high, people blame shortcomings on factors that may be unrelated to the actual problem (Kruger & Dunning, 1999). If feedback is vague or unclear, leaders are less likely to take ownership of the problem or even acknowledge that a problem exists at all.

Summary of measurement. Two types of response formats are typically used to measure competencies on multi-rater instruments: frequency scales and evaluative scales. Frequency scales help identify shortcomings, or deficits in a leader's behavior. Evaluative scales help to identify areas of strength. Each scale is afflicted with a blind spot, however. Frequency scales are confounded at the higher end of their register. They are unable to distinguish between the optimal amount of a behavior and too much of a behavior. Therefore "high" ratings could mask problematic behaviors. Evaluative scales are confounded at the lower end of their register. Respondents are unable to provide feedback on why the leader rated poorly. Due to these blind spots, measuring competencies should not be mistaken with measuring competent leadership. Effective leadership is a complex phenomenon where leaders assess the situational context, the needs of the people involved, and leaders apply the correct skill to fit that situation (Kaplan & Kaiser, 2006). Competency models risk oversimplifying this complexity and missing their mark (Hollenbeck & McCall, 2006).

The Versatile Leader Model

Leadership Versatility

The versatile leader model was designed to help executives develop a well-rounded approach to leadership (Kaplan & Kaiser, 2006). Kaplan recognized that many executives with whom he consulted received consistently high 360-degree instrument ratings, but co-workers' narrative comments suggested problems not reflected in quantitative feedback (Kaplan & Kaiser, 2006). Subsequent observation, research, and analysis revealed that leadership instruments were not sensitive to overuse of strengths. First Kaplan (1996) and then Kaplan and Kaiser (2003; 2006) developed the versatile leader model and a multi-rater feedback instrument, the Leadership Versatility Index (LVI), to address this shortcoming.

Kaplan (1996) observed that the root of the word versatility meant “to turn around” or “to pivot” (p.1). He conceptualized versatility as the ability to pivot from one style or approach to another depending on the needs of a situation. Kaplan and Kaiser (2003; 2006) presented leadership versatility as a dynamic relationship between seemingly opposite virtues. Sometimes leaders are forceful, imposing their will to accomplish objectives; other times they enable others to take the reins and chart a course of action. Leaders may envision future strategies and encourage innovative projects or they may emphasize day-to-day operational efficiency. Depending on situational needs, interpersonal behaviors and leadership initiatives have benefits and consequences.

Kaplan's model emphasized the dynamic balance among behaviors (1996). Leaders who struggled to move freely between behaviors were considered lopsided, or

biased toward one side of a duality or the other. Not only were lopsided managers biased, but the greater their lopsidedness, the more prejudiced they were against its complement (2003). To illustrate, a manager heavily biased toward forceful leadership will not only prefer directive actions, but will also scoff at using inclusive, enabling behaviors.

Enabling behaviors will be characterized being overly soft and ineffective (Kaplan & Kaiser, 2006). These findings were consistent with those reported by Vroom and Yetton (1973)—managers restricted their decision-making criteria toward personal biases and preferences. Versatile leaders, in contrast, moved fluidly between dualities (Kaplan & Kaiser, 2003). They used the appropriate behavior in a timely fashion without depending on too narrow a range of options (Kaplan & Kaiser, 2006). Expert leaders may intuitively grasp how to balance strengths without overusing them, but the versatile leader model organizes leadership behaviors such that the dynamic balance of strengths and weaknesses can be empirically examined. According to Kaplan and Kaiser (2003), these expert leaders are far and few between—less than one in five managers would qualify as versatile.

The impact of versatility is powerful. In a sample of nearly 700 executives, the correlation between high versatility and overall effectiveness was .71 (Kaiser & Kaplan, 2006). Said another way, versatility accounted for half of the variance in leader effectiveness. In a recent analyses of data produced by a refined set of items, versatility accounted for more than half of the variance for effectiveness (Kaiser, 2009, personal communication). Possessing an array of competency-based leadership skills is important, Kaplan and Kaiser's (2006) research indicated no single quality was more important than

versatility. Indeed, professional excellence is not defined by an individual's knowledge or skills, but rather by the practitioner's ability to apply them meaningfully (Schon, 1983).

Conceptual structure of the Leadership Versatility Index

Traditional instruments measure leadership behaviors as isolated, unrelated aptitudes (McCall, 2006). They cannot measure whether behaviors are dynamically balanced (Kaplan & Kaiser, 2003). Balance is central to the concept of versatility and to its measurement practices (Kaplan & Kaiser, 2003; 2006) Balance is reflected in the LVI through two design features: the Too Little/Too Much scale and the duality-based structure (Kaplan & Kaiser, 2003).

Too Little/Too Much response scale. Kaplan and Kaiser (2003; 2006) incorporated a unique response scale into the LVI called the Too Little/Too Much scale. It's a bi-directional scale with values ranging from -4 to +4 and the ideal score is 0, the central position. When leaders use a behavior effectively, raters score them with 0 (*the right amount*). If leaders underutilize a behavior, raters indicate the degree of under-use by scoring between -1 (*barely too little*) and -4 (*much too little*). Conversely, when leaders overuse a behavior, they are scored between +1 (*barely too much*) and +4 (*much too much*). This approach allows observers to simultaneously provide judgment on the effectiveness of a behavior and also provide feedback on frequency. For this reason, the Too Little/Too Much response scale may be considered an *evaluation-of-frequency* scale. In effect, each item provides its own frequency and effectiveness bell curve (Kaiser & Kaplan, 2005). Ratings scored in the middle of the response scale are those a rater considers most effective; and ratings that diverge from center are perceived to be less

optimal. Depending on the direction of the bias, leaders can discern whether they are overusing strengths or are deficient (Kaiser & Kaplan, 2005).

This response scale mirrors principles within Aristotle's Doctrine of the Mean. Aristotle (trans. 2004) taught ethical virtue as a condition that rests between opposing states of excess and deficiency. He stipulated that this mean was a fluctuating target that was dependent on circumstances, "[T]he mean is to be determined in a way that takes into account the particular circumstances of the individual ... Finding the mean in any given situation is not a mechanical or thoughtless procedure, but requires a full and detailed acquaintance with the circumstances" (Kraut, 2008, Ethical Virtue as Disposition, para. 2). There is no fixed measure for ethical behavior; rather, ethical virtue is a product of experience, effective guidance, and wisdom. Given the contingencies for which leaders must account, leadership effectiveness also fluctuates on a fulcrum between excess and deficiency (Kaiser & Kaplan, 2005). Traditional frequency and evaluation response scales are unable to reflect this balance because they cannot measure excess (Kaiser & Kaplan, 2005).

Duality-based structure. Balance is also reflected through the duality-based structure of the LVI. To measure dynamic balance, Kaplan and Kaiser (2006) arranged leadership behaviors into sets of paradoxical relationships called dualities. A duality is composed of two opposed, but significantly related dimensions. Chinese philosophers conceived the *yin* and *yang* duality to explain a universal principle of how opposing forces complement and consume one another and thus impact the surrounding environment (Billington, 1997). One side of a duality is the necessary complement to its

mate (Billington, 1997). The leader who always talks and never *Listens* does not truly communicate. Conversely, the leader who always *Listens* and never talks is also a poor communicator. Both *yin* and *yang* are necessary.

Two dualities were defined. Leaders influence constituents through interpersonal processes—either through forceful behaviors or enabling behaviors. The forceful dimension addressed leaders means of asserting their own power, capability, and authority (Kaplan & Kaiser, 2003a). The enabling dimension was about creating conditions for others to lead (Kaplan & Kaiser, 2003a). Therefore the forceful/enabling duality focused on *how* leaders influence others through their leadership. Leaders also drive results through initiatives—either through strategic vision or through operational efficiency. The strategic dimension focused on long-term strategy and external focus. The operational dimension focused on short-term execution and an internal focus. Thus, the strategic/operational duality focused on *what* initiatives leaders choose to lead (Kaplan & Kaiser, 2006). The LVI was crafted with parallel structure, where the design one side of a duality was mirrored on the other. Each duality was divided into two dimensions. These dimensions were further divided into three sub-dimensions and each sub-dimension was composed of four items (Kaplan & Kaiser, 2006). To achieve balanced properties, items are arranged in matched pairs. At the item level each item has a complementary mate in a corresponding sub-dimension of the opposing duality. Taken as a whole, the duality is composed of a series of matched pairs at the item level, the sub-dimensional level, and the dimensional level (Kaplan & Kaiser, 2006).

Forceful / Enabling duality. In his seminal review of leadership literature, Bass (1990) organized theories into an overarching cluster entitled “autocratic and authoritarian leadership versus democratic and egalitarian leadership,” often referred to as autocratic versus democratic leadership (Bass, 1990, p. 415). This superordinate category encompassed nearly thirty dichotomy-based concepts published between 1938 and 1985 (Bass, 1990). These theories were primarily of the behavioral and contingency type, and most explored the nature of task-focused leadership versus relationship-focused leadership. Research conducted by Edwards and Rode (1986) studied these theories to determine whether the constructs under scrutiny were differentially unique (Bass, 1990). Their findings showed positive but lower than expected correlations between selected autocratic and task-centered constructs; the same positive, but lower than expected result was found with democratic, relationship-centered constructs (Bass, 1990). It appeared that although constructs within each side of the autocratic-democratic duality had thematic similarity, they were distinctive enough to avoid redundancy.

Kaplan and Kaiser (2006) based the forceful and enabling duality on the autocratic and democratic framework articulated by Bass (1990). The forceful dimension was analogous to autocratic leadership and the enabling dimension was analogous to democratic leadership. Leaders can act forcefully in different ways just as leaders can encourage others to lead through various methods. Therefore, three theoretically distinct perspectives subdivided the forceful and enabling dimensions: power, decision-making styles, and orientation (Kaplan & Kaiser, 2006). These three perspectives will be

discussed as well as the three forceful and three enabling sub-dimensions that they help define.

Perspective one: Locus of power. The first perspective, and the pair of sub-dimensions derived from it, address locus of power. Leaders can wield power in various ways; for example, sometimes a leader steps in and commands the till. *Takes Charge* represents a forceful, self-driven locus of power. At other times a leader might imbue others with the opportunity to chart a course. *Empowers others* represents an enabling locus of power focused on followers. The locus of power perspective has roots in Stogdill & Coons' (1957) work on initiating structure (forceful) and consideration (enabling). It also owes lineage to Likert's (1967) influence skills (forceful) versus interaction skills (enabling) and Zaleznik's (1974) power-oriented (forceful) versus power-sharing (enabling) leadership models. The Leadership Versatility Index has four items that measure *Takes Charge* and four that measure *Empowers others*. Just as the dimension and sub-dimensions have complementary opposites of one another, each item has its complementary opposite too. For example, item 1f in *Takes Charge* is the complementary opposite of Item 1e in *Empowers others*.

Perspective two: Decision-making. The second pair of sub-dimensions addresses decision-making. Leaders can choose to make decisions autocratically or they can make them democratically. *Declares* represents a forceful decision-making processes centered on a leader's independent judgment whereas *Listens* represents a more inclusive and participative decision-making process. Lewin and Lippett's (1938) autocracy and democracy concepts, House's (1971) directive versus participative behaviors, Vroom and

Yetton's (1973) directive and autonomous decision making versus participative and inclusive decision-making, and Bass and Valenzi's (1974) directive/persuasive approach versus consultative/participative approach represent historical precedents for the decision-making perspective. As in the previous case, four matched pairs of items measure *Declares* and *Listens*.

Perspective three: Orientation. The third pair of sub-dimensions addresses motivational orientation. *Pushes for performance* represents a forceful orientation with a primary concern for task completion whereas *Supports* represents an enabling orientation with primary concern for people and positive relationships. Numerous models have framed leaders' motivational orientations. Bales (1950) articulated orientation in terms of performance (forceful) versus maintenance (enabling) behaviors. Blake and Mouton' (1964) managerial grid incorporated orientation within the production emphasis versus people emphasis constructs, as did Fiedler (1967) in the task-oriented versus people-oriented aspect of contingency theory. Hersey and Blanchard (1969) offered similar concepts in the form of direction and support, House's path-goal theory addressed achievement-focused behavior versus supportive behavior, and Quinn (1988) addressed motivational orientation in the producer and director roles versus the mentor and facilitator roles. Four matched item pairs measure *Pushes for performance* and *Supports*.

Even though pairs of sub-dimensions are conceptually related to one another, they are not polar opposites of a single underlying continuum. Instead they are related but empirically distinct constructs (Kaplan & Kaiser, 2006). In summary, the forceful/enabling duality is comprised of six distinct sub-dimensions, linked through pair-

wise relationships. The empirical relationship between forceful/enabling sub-dimensions and the overall factor structure will be discussed in a section presenting the psychometric properties of the LVI.

Strategic / Operational duality. The strategic/operational duality addressed elements of organizational direction and execution (Kaplan & Kaiser, 2006). This duality can be of great importance where executives and higher-level middle managers are concerned because as job complexity increases, so does the need to think and function on behalf of the organization (Kaiser, Craig, Overfield, & Yarborough 2009). These dimensions are not relevant to the management levels under consideration in this study (hall directors and resident advisors), and they will not be included in the LVI-S. They will be briefly reviewed, but in limited detail.

The strategic/operational duality focuses on what type of business or organizational matters upon which leaders focus their time and attention (Kaplan & Kaiser, 2006). Kotter (1990) juxtaposed the differences between leading and managing. Leaders established visionary direction for their organizations; managers planned how to execute these visions. Leaders were change agents who sometimes eschewed structure, whereas managers preserved structure and control. The strategic dimension and operational dimensions differ along similar lines. Kotter's leadership construct was similar to the strategic dimension and the management construct was similar to the operational dimension (Kaplan & Kaiser, 2006).

Like forceful/enabling dimensions, strategic/operational dimensions are also divided into three pairs of sub-dimensions. The first pair of sub-dimensions addresses

timeline and activity. *Long-term direction* represents strategic emphasis whereas *short-term execution* represents operational emphasis. The second pair of sub-dimensions addresses orientation. *Growth* represents a strategic concern with expanding capability and capacity, whereas *efficiency* represents an operational concern about conserving resources and focusing effort. The third pair of sub-dimensions addresses climate. *Innovation* represents the extent to which leaders encourage creativity and support for employees trying new approaches, whereas *order* represents a more systematic process governed by disciplined procedures (Kaplan & Kaiser, 2006).

Psychometric Properties of the Leadership Versatility Index

The Forceful / Enabling factorial structure. The Forceful/Enabling duality, dimensions, and sub-dimensions of the LVI have been empirically supported through a series of structural factor analyses (Kaiser & Craig, 2001; Kaplan & Kaiser, 2006). Kaiser and Craig (2001) explored three possible models for the structure of the forceful-enabling duality: (1) a one-factor model where forceful leadership corresponded to one end of the continuum and enabling leadership on the other; (2) a two-factor model where forceful and enabling leadership were uncorrelated; and (3) a two-factor model where both factors were inversely related, indicating that when one factor is overused, the other will be underused. Seventeen items gleaned from an exploratory factor analysis were used to measure the three forceful sub-dimensions (*Takes Charge*, *Declares*, and *Pushes*) and the three enabling sub-dimensions (*Empowers*, *Listens*, and *Supports*). Items loaded on their intended sub-dimensions, and these sub-dimensions loaded on their respective forceful or enabling dimension. The forceful and enabling dimensions were inversely

correlated ($r = -.58$, corrected for measurement error) (Kaplan & Kaiser, 2006).

Therefore, the two-factor, inversely related model provided the best fit for the data (Kaiser & Craig, 2001; Kaplan & Kaiser, 2006). These findings supported the duality effect—as leaders emphasized one side of a duality, they under-used the other side (Kaplan & Kaiser, 2006). As Chinese philosophers observed, opposing forces both complement and consume one another, thereby impacting the surrounding environment (Billington, 1997).

Reliability of the LVI. Reliability for the LVI was considered in two forms. First, reliability was reported at the individual level of analysis. Chronbach's alpha provided a measure of internal consistency. Nunnally (1978) suggested a rating of .70 as the minimum acceptable standard, but .80 is preferred for practice. The forceful and enabling scales both exceeded .80 across all three rater categories (superiors, peers, and subordinates). The alpha for the strategic scale was .79, an acceptable score, but the operational scale required more work as its alpha was only .53 (Kaplan & Kaiser, 2006). Reliability for multi-rater instruments also considers the degree of convergence between multiple raters of the same target (Kaplan & Kaiser, 2006). Kaplan and Kaiser (2006) followed the recommended procedures of LeBreton, Burgess, Kaiser, Atchley, and James (2003), providing statistics for both inter-rater reliability and inter-rater agreement.

Inter-rater reliability provides an index of the extent to which ratings of the same target, made by different observers, are consistent (follow the same path of highs and lows) (Kaplan & Kaiser, 2006). Interclass correlation coefficient (ICC; Shrout & Fliese, 1976) was used to calculate inter-rater reliability (Kaplan & Kaiser, 2006). These ratings

were calculated at the single rater level ICC(1) as well as at the rating group level [ICC(k), where k = number of raters]. Kaplan and Kaiser (2006) stated the greater the number of raters, the greater the reliability. ICC(1) values should range from .2 to .45 and ICC(k) values should exceed .70 although values between .50 and .70 are usually considered acceptable (Kaplan & Kaiser, 2006). Values across all forceful/enabling scales were in the range of acceptability.

Inter-rater agreement was determined using the r_{wg} statistic (Kaplan & Kaiser, 2006). This provides an indication of the degree to which different raters for the same target issued the same rating on each item (Kaplan & Kaiser, 2006). Mean r_{wg} values greater than .70 are regarded as acceptable in most cases and values approaching 1.00 indicate very high agreement (Bliese, 2000). The mean r_{wg} values met or exceeded .85 across all categories (Kaplan & Kaiser, 2006).

Validity of the LVI. Kaplan and Kaiser (2006) reported that two concerns should be considered when considering the validity of the LVI. The LVI purports to be a duality-based instrument. If the model operates as intended, then there should be a negative correlation among opposite behaviors. Kaplan and Kaiser (2006) refer to this as a “polarity effect” (p. 201). It is the statistical representation of a manager’s tendency to gravitate to one set of behaviors and exclude their complement (Kaplan & Kaiser, 2006). The strategic and operational dimensions did not exhibit the polarity effect, but the forceful and enabling dimensions provided average correlations among groups between -.56 and -.61 (Kaplan & Kaiser, 2006). The second consideration is how well the data fit the conceptual structure. The exploratory and confirmatory factor analyses provided an

adequate measurement model for the structure of the forceful and enabling dimensions (Kaiser & Craig, 2001; Kaplan & Kaiser, 2006).

Value of LVI as a Development Tool

Competency models are used to train and develop leaders (Lombardo & Eichinger, 2000). Advocates of competency models claim they help specify a range of useful leader behaviors, provide a tool for personal career and leadership development, and outline a leadership framework that can be used to develop, select, and understand leadership effectiveness (Silzer, 2006). Opponents claim that leadership is contextual and highly complex. Competency models oversimplify the nature of this challenge (Hollenbeck & McCall, 2006). The LVI helps build a bridge between these two camps. The LVI has a duality-based structure that more accurately reflects the tensions and tradeoffs characteristic of leadership. It provides a method for empirically assessing the biases inherent in people's leadership styles. It also permits leaders to more effectively grasp the interdependent nature of behaviors and the implications of choosing the familiar option over the foreign option. The LVI provides item-specific feedback due to the response scale being a judgment of frequency scale. Each item, in effect, has its own bell-curve for effectiveness. Optimum effectiveness rests at the center of the scale, and as the absolute value increases from zero, leaders know they are either deficient or are overusing strengths (Kaiser & Kaplan, 2005). Such feedback can help them become more versatile in their leadership (Kaiser & Kaplan, 2006).

Leadership is a paradoxical phenomenon (Martin, 2007). An assessment that reflects paradox provides greater insight into the nature of these challenges and the

psychosocial tension they place on leaders (Kaplan & Kaiser, 2006). Leaders can receive immediate, unambiguous feedback that helps them understand more about their shortcomings, strengths, and strengths overused (Kaiser & Kaplan, 2003; 2006). These insights are especially valuable in cases where leaders must learn to negotiate complex issues and do so with a measure of interpersonal savvy.

Some have argued that leadership development is at an important crossroads for our culture (Drucker, 1998; Kaiser, 2005; Tierney, 2006). Lombardo and Eichinger (2000) noted that organizations are going to experience a leadership supply problem:

The demand for those who can deal with change and ambiguity and fragmentation is increasing, but the supply of people who can do those things is pretty much what it has always been. (And according to McKinsey, that supply will shrink over the next 15 years). In AT&T's assessments of high potential managers, only 12 percent could cope effectively with the ambiguity and fragmentation of rapidly changing conditions. So the need for leaders is greater than ever, but there is a problem on the supply side. (p. 2)

The importance of leadership development is as critical as it has ever been. Programs and processes need to evolve to keep up with the increased demand for more leaders and more capable leaders. These leaders need to be coached in paradox and the nature of complex thinking (Martin, 2007).

One final observation about the LVI and its potential impact on leadership development: Executives are a small population with great influence, but some might argue they have a wealth of development resources from which to choose. The LVI was designed for these executives, but a pending demographic shift creates a demand for new tools to develop leaders at an earlier stage of their careers. This study is focused on

designing and validating the LVI for a college campus-based population. Leadership development is an iterative process (McCauley & Van Velsor (2004). The more one learns the more one is able to learn. If students and student-affairs professionals can learn more sophisticated techniques earlier in their careers, they have the potential to develop into better leaders, faster.

“The Future That Has Already Happened”

Drucker (1998) alleged that predicting the future was a fruitless enterprise; but he also stated that in the course of societal life, major events occur that have predictable, consequential impacts over several decades. From these events, Drucker (1998) extrapolated “the future that has already happened” (p. vii). Since the latter part of the twentieth century, demographics have been, and will continue to be, the primary driver of social and economic change in the developed world (Drucker, 1998). Developed countries have slowly been sowing the seeds of self-destruction because its citizens were not producing enough babies to reproduce themselves (Drucker, 1998). Drucker (1998) stated the impacts of population decline will be vast: The retirement age for healthy workers will increase to 75 before 2010; economic growth will no longer come from increasing the workforce nor through greater consumer demand, but only through the productivity of knowledge work and of knowledge workers; the only way to maintain competitive standing in the world economy will be through ongoing, systematic improvements to the production of knowledge and the productivity of knowledge workers; the global economy will be highly turbulent and interconnected; the information

needs of executives and businesspeople will shift greatly; and finally, organizations and workers roles within organizations will dramatically change.

Through the latter half of the 20th century, citizens of developed nations produced fewer children; consequently, the demographic profile of developed nations has grown increasingly older. Costs associated with supporting an aging population necessitated younger families cut back on child dependents (Drucker, 1998). Developed countries were unable to sustain economic growth through putting more people to work due to the shortage of workers available to service low paying production jobs (Drucker, 1998). Subsequently, these jobs have been outsourced to countries better suited to manufacturing—countries whose economies are in an emergent stage and countries that have an abundant supply of workers (e.g., China, India, etc.) (Friedman, 2007; Drucker, 1998).

Developed countries adapted to these circumstances, managing to expand their economies through efficiency, productivity, and effective leverage of knowledge (Drucker, 1998). In the latter half of the twentieth century, economically developed countries such as the United States and those in Western Europe shifted economic emphasis from production of goods and services to production of knowledge (Drucker, 1968). Malchup (1962) identified this trend as the emergence of the “knowledge economy.” Since workers were not abundant, organizations learned how to do more work with fewer people. Ultimately knowledge became a product in-and-of itself. Jobs and services associated with knowledge-work were higher paying and more stable (Drucker, 1968). Industries dedicated themselves to process efficiency, created technologies that

enhanced efficiency, and provided services that networked knowledge centers (Friedman, 2007).

Shifting Leadership Paradigms

The knowledge-based economic engine necessitated a different approach to leadership and management. Organizations were changing from hierarchical, command-and-control structures characteristic of the industrial era into flat, interdependent networks more suited for the information age. Drucker (1998) illustrated the nature of these changes, “As more and more organizations become information-based, they are transforming themselves into ... responsibility-based organizations in which every member must act as a responsible decision maker. All members, in other words, have to see themselves as ‘executives’” (p.126).

A shift in leadership paradigms accompanied this change in organizational management. Leadership models became more process-focused, meaning leadership was the product of social interactions at all levels within an organization—not simply the actions of a person with positional authority or power (Kouzes & Posner, 2002; Rost, 1993). Northouse (2004) referred to this “emergent leader” as one who has not been assigned a leadership role, but who influences others through interpersonal communication skills, seeking others opinions, initiating ideas, and being able to take a stand without appearing stubborn or rigid (p. 6). According to Drucker (1998), workers were no longer considered a necessary expense required for the production of goods. In the information age, workers are considered a critical input for the knowledge generating process (Drucker, 1998).

Drucker (1998) cautioned that the most difficult challenge that developed countries will face is increasing the productivity of knowledge workers—especially in service sectors:

The chief *economic* priority for developed countries ... must be to raise the productivity of knowledge and service work ... The most pressing *social* challenge developed countries face, however, will be to raise the productivity of service work. Unless this challenge is met, the developed world will face increasing social tensions, increasing polarization, increasing radicalization, possibly even class war. (p. 144, italics in orig.)

These observations precisely forecast the challenges our society faces—as indicated by Astin and Astin (2000), Tierney (2006), Putnam (2000), a national study of social capital (Easterling & Foy, 2006) and aforementioned current events. Efforts to teach leadership and develop leaders must keep pace with both economic and societal leadership demands. For this reason, colleges and universities serve a critical role in preparing students for the leadership demands of the foreseeable future.

Development of Leadership in College and University Contexts

University mission statements proclaim the importance of developing leaders, providing service to community, and preparing students for the world of tomorrow (Morphew & Hartley, 2006; Smart, Ethington, Riggs, & Thompson, 2002). Astin and Astin (2000) proposed that colleges were strategically suited to address the pending leadership demands and Higher Education could serve a critical role in developing leaders equipped to address challenges of the present and future. College has multiple contexts where leadership can be practiced and learned—through academic classes, certificate-based leadership programs, student government, and student employment,

among others (Eich, 2007). Smart, et al., (2002) conducted a path analysis of institutional spending patterns to better understand which leadership-related expenditures influenced students' perception of institutional commitment to leadership development. The study reported that extracurricular programming was a significant contributor to student leadership growth (Smart, et al., 2002).

Smart, et al. (2002) utilized path analysis to explore the relationship between university expenditure patterns and students' self-reported leadership growth. The path analytic technique permitted researchers to examine the direct and indirect effects of expenditure patterns and distinguish among different types of expenditure categories (Smart, et al., 2002). Their findings indicated a negative relationship between instructional expenditures and student reports of growth. The authors asserted this negative relationship indicated that leadership development responsibility is not shared across equally across academic disciplines. More enterprising disciplines enhanced students' perception of leadership growth although others did not (Smart, et al., 2002). The reverse was true for student service expenditures (student affairs expenditures) and self-reported growth. Smart, et al., (2002) reported a significant positive relationship between student service expenditures and the likelihood that students will participate in leadership activities during their undergraduate careers. These students also perceived their institutions as placing a priority on student leadership development (Smart, et al., 2002). Said another way, extracurricular activities were a critical factor in student leadership growth and development.

How Do Leaders Develop in College and University Settings?

Numerous studies have reported that the college experience helps students increase their leadership ability, but identifying a specific setting or context that enhances student leadership development is quite difficult because of the range of interactions college life offers (Pascarella & Terenzini, 2005). Pascarella and Terenzini (2005) speculated that gains in interpersonal skills appeared to result from the cumulative effect of interpersonal contacts and relationships rather than to a specific context or structure for such contacts. Students develop leadership skills through involvement with leadership positions such as student government or student employment (Pacarella & Terenzini, 2005; Posner & Rosenberger, 1997; Smart, Ethington, et al., 2002), participating in student organizations or volunteering where peers interact with other peers (Astin, 1993; Astin & Sax, 1998; Pike, 2000), ROTC (Baxter, 2001), leadership programs and classes (Cress, Astin, Zimmerman-Oster, & Burkhardt, 2001; Eich, 2008; Zimmerman-Oster & Burkhardt, 1999a; 1999b) and general activities where students engage in leadership responsibilities (Pascarella & Terenzini, 2005).

Pascarella and Terenzini's (2005) speculation that leadership gains were due to cumulative effects rather than to a specific form or setting could be premature, however. Research methodology could lie at the heart of this problem. The majority of recent student leadership research has not studied how well student leaders have acted as leaders; rather, it has focused on the environmental effects college has had on the leader. An evaluation of student leadership literature by Pascarella and Terenzini (2005) evidenced a shift away from behavioral, interpersonal leadership research toward studies

exploring the environmental impacts college has on student leaders (e.g., Dugan & Komives, 2007; Astin, 1993). Pascarella and Terenzini (2005) observed that prior to 1990, the emphasis of study relied heavily psychological constructs and measures but this focus shifted dramatically after 1990:

[T]he post-1990 research has largely ignored interpersonal relations in favor of examining the influences of institutional characteristics on students' self-reported abilities in areas related to leadership. These skills are often measured by composite scales reflecting traits such as self-confidence, ability to get along with others, popularity, or leadership positions held ... the variables generally used to differentiate among institutions appear to have little predictive value. (p. 236)

Concerns with Student Affairs Leadership Literature

Methodological Concerns

Pascarella and Terenzini's (2005) observations bring several concerns to light, none of which appears to have been acknowledged in student affairs leadership literature. First, this body of literature is highly dependent on self-report surveys, but few analytical studies in this literature have questioned whether student self-reports are a valid source for leadership data (Turrentine, 2001). One fundamental problem with self-reported research, and a likely reason there is little convergence between self-report data and other-reported data, is because multi-rater studies have demonstrated that bad leaders believe they are good and good leaders believe they are average (Atwater & Yammarino, 1992; Conway & Huffcutt, 1997; Harris, & Schaubroeck, 1988). In general, leaders are poor estimators of their ability. Leadership is a social behavior (Denison, Hooijberg, & Quinn, 1995); and as such, leaders engage in observable actions that connote effective or

ineffective leadership. Observer reports of these actions are reliable, valid predictors, whereas a leader's self-reported beliefs or feelings about what they thought they did or intended to do are not (Atwater & Yammarino, 1992; Conway & Huffcutt, 1997; Harris, & Schaubroeck, 1988).

Second, the level of analysis is a critical factor for researchers to consider. Yammarino, Dionne, Chun, & Dansereau (2005) provided a qualitative review of leadership literature that focused on levels of analysis research. They identified four entities, or levels, at which analysis can be conducted: individual, dyadic, group-level—interdependent groups of people who interact with one another, and collectives (e.g., groups of departments who might be part of a functional team but do not directly interact, a hierarchical institution). The individual or small-group leadership research agenda and institutional research agenda can be largely separate exercises. Postulating that findings from one source inform the practice of the other is a flawed assumption known as an ecological fallacy (Antonakis, Schriesheim, Donovan, Gopalkrishna-Pillai, Pellegrini, & Rossome, 2004). Each agenda investigates leadership at a different level. Research has demonstrated that significant correlates at one level of analysis may or may not be significant at other levels (Antonakis, et al., 2004; Atwater & Yammarino, 1992; Yammarino, Dionne, Chu, and Danserau, 2005). Cross-level and multi-level research designs have been employed, so searching for meaningful findings at the individual and institutional levels is not a fruitless endeavor (Yammarino, et al., 2005). But scholars noted that research designs need to account for proper measurement of constructs and use data analytic techniques that correspond to the intended level of analysis so that

inferences are not misleading or produced by statistical artifacts (Yammarino, et al., 2005). Multi-rater assessment methods and research designs incorporating levels of analysis into their analytical procedures can complement the existing body of institutionally focused student leadership literature to provide a more comprehensive network of leadership knowledge.

The first section addresses the validity of self-report data. It begins with a critical review of the only known student-affairs study that compared the congruence of self-report data to peer-report data (Turrentine, 2001). Since self-report data occupies such a prominent position in student-affairs leadership research, a meta-analysis on the conditions of validity necessary for self-report data is also reviewed (Mabe & West, 1982). Podsakoff and Organ (1986) reviewed the problems and prospects associated with self-report collection methods. Podsakoff and Organ's (1986) insights are relevant to the critical analysis of leadership survey data—especially where the use of criterion variables and predictive validity are concerned.

The second section compares self-report data to observer report data. Two meta-analyses are reviewed, the first analyzed correlations between self raters, peer raters, and superior raters (Harris & Schaubroeck, 1988) and the second study extended the self-observer correlation analysis to include subordinate ratings (Conway & Huffcutt, 1997). Findings indicated frequent discrepancies between self-raters and observers. Since these discrepancies could involve self-inflated ratings (Harris & Schaubroeck, 1988; Conway & Huffcutt, 1997) or self-deflated ratings (Conway & Huffcutt, 1997), a study on the directional implications of self-ratings is reviewed (Kruger & Dunning, 1999). The final

category addresses levels of analysis. Levels of analysis are explained and the implications of not accounting for levels of analysis are stated. Atwater and Yammarino (1992) provide an example of a multi-rater, multi-level study that provided significant results at one level, but at the overall, aggregate level these results were obfuscated. Research designs that do not account for levels of analysis put researchers at risk of interpreting incomplete or bad data, or missing significant results hidden by aggregate statistics.

Critical Analysis of Self-report Validity and Student Affairs Research Methods

Turrentine (2001) observed, “Much of the research in student affairs relies on student self-reports of their own behavior...student affairs scholarship and practice depend on the largely unexamined assumption that students’ self-reports are both honest and accurate” (p. 361). Student affairs professionals have incorporated self-report designs because they are fast, less expensive, and easy to administer. Turrentine’s study tested whether self-ratings and peer ratings of leadership skills were congruent with one another. It was the only study found in student affairs literature that expressly studied the validity of self-ratings compared to other-reports. The population consisted of 108 undergraduate members of a residential leadership community at a large, Southeastern public university in the United States (55% female, 45% male, 85% were first-year students). Turrentine’s overall analysis indicated there was a significant difference between self-reports and observer reports ($t=3.163$, $df = 107$, $p = .002$). Student self-reports, on average, were more favorable than peer reports. Turrentine attributed these differences to the self-rater’s ability to observe all behaviors, whereas peers only

observed a portion of the rater's actions, hence raters credited themselves with all of their leadership actions whereas peers could only credit them with those they observed. She noted Kouzes and Posner (1988) reported similar findings and accounted for their findings through a similar rationale. Though not acknowledged by Turrentine, discrepancies between self and observer ratings have been reported elsewhere in multi-rater meta-analyses (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988) as well as numerous multi-rater studies conducted with student populations (Atwater & Yammarino, 1992) or in student affairs settings (Komives, 1991a; 1991b; 1991c; Posner & Brodsky, 1993). The limited observation hypothesis has been intimated in these other studies, however additional reasons have been provided for inflated self-ratings. A more detailed discussion of alternative hypotheses is warranted and will be provided when reviewing the studies noted above.

Turrentine's (2001) discussion of results was curious because she appeared to ignore the observed findings. Even though data indicated a significant difference between self- and peer ratings, she stated:

It is highly likely that students were both accurate and honest in claiming to have practiced skills that the peer reports could not have confirmed. Nevertheless, more than seven in ten of the self-reports of leadership behaviors were confirmed by peer observations... in the great majority of cases ... the peer observations *were* congruent with those supplied by the student leaders about themselves (pp. 369-370; italics in original).

In the summary of implications, Turrentine (2001) proposed:

This study began with the observation that student affairs scholarship and practice rely on student self-reports. The results of this study, if confirmed in future research, provide a basis for confidence in students' accounts of their own behaviors. (p. 371 italics added)

In the methodology section of a recent leadership study that investigated self-efficacy in commuter students, Dugan, Garland, et al., (2008) reported Turrentine's (2001) study as evidentiary support for self-rater methodology, "a study on self-and peer reported leadership behaviors and the quality of those behaviors found self-reports of leadership to be generally accurate" (p. 288). Turrentine's data did not support this conclusion, but interpretive comments within the discussion clouded these findings. The Turrentine discussion appeared to provide support for self-report methods in leadership research, but this claim was not supported by data.

A meta-analysis of self-report data. Mabe and West (1982) reviewed literature pertaining to the validity of self-evaluation of ability. They conducted a meta-analysis of 55 studies where ability evaluations were compared with measurements of performance. A total of 14,811 participants were included in these studies, with 81% of them being college students. The studies incorporated 103 criterion measures, of which objective tests (51 examples), class grades (19 examples), and supervisor ratings (16 examples) were most frequently identified as criteria. A total of 267 correlation coefficients were examined. According to Mabe and West's literature review, self-evaluation of ability may closely correspond to performance on criterion measures provided that certain measurement conditions have been addressed. Although nine measurement conditions

were identified, the study found four to be significant. In order of predictive significance, these were:

1. When participants expect their self-evaluations will be compared with actual criterion measures, the validity of their responses improved.
2. Evaluations should provide a meaningful rating context; they need to provide information about one's relative standing with respect to a social reference group. Evaluation of ability is a more of a relative process than an absolute process. Provided a referent has been specified, using response prompts like below average, average, and above average connote relative ability whereas prompts such as poor, fair, and excellent are more absolute and thus less valid.
3. The more experienced a participant is at self-evaluation (with respect to the criterion being measured), the more valid their responses are. This improvement probably reflects greater self-knowledge with regard to past performance and an increased ability in self-assessment.
4. Anonymity of responses increases validity.

Overall, the mean validity coefficient for the studies reviewed by Mabe and West (1982) was a low, unweighted correlation coefficient of .29. There also was a high degree of variability ($SD = .25$). Of the 267 correlations Mabe and West considered within the study, 203 (76%) were obtained under conditions where only three or fewer of the nine favorable measurement conditions were met. There was a general increase in mean validity as the number of favorable conditions increased, ranging from .00 when no conditions were met ($SD = .14$) to .63 when all four were met ($SD = .17$). The limited

number of correlations in this most favorable group (4), however, indicated that these results should be interpreted with caution.

Mabe and West (1982) provided information that serves dual purposes. They provided standards that may be used to evaluate the efficacy of self-report research designs. Given the prominence of self-report data in student affairs literature, knowledge of and implementation of these guidelines can be beneficial to increase validity of self-report data. But these guidelines also serve to inform multi-rater assessment as well. Criterion variables should be used to enhance validity, group-specific (social comparison) terminology should provide participants a frame of reference, and anonymity contributes to instrument validity. The importance of rater experience may be the most noteworthy point. Mabe and West (1982, p.294) stated:

In the self-evaluation studies reviewed, the provision of self-evaluation experience was more incidental than systematic, and it may be that more careful structuring of these self-evaluation experiences may help elicit valid self-evaluation... It seems reasonable to hypothesize that conditions increasing objective self-awareness would also operate to increase the validity of self-evaluation.

Problems with self-report leadership data. Given the importance of using criterion variables to increase the validity of self-report assessment (Mabe & West, 1982), Podsakoff and Organ (1986) provided cautionary advice relevant to measuring self-reported leadership skills and effectiveness ratings through questionnaire-based methods. Although they did not discourage the practice, these researchers recommended several procedures that should be used to control for artificially high correlations. Podsakoff and Organ (1986) stated that self reports have been used to gather data for

approximately six different purposes: (1) Obtaining demographic or factual data; (2) assessing effectiveness of experimental manipulations; (3) gathering personality data; (4) gathering descriptions of a respondent's past or characteristic behavior or how the respondent might have behaved under specific hypothetical conditions; (5) scaling respondents' psychological states (e.g., job attitude, tension, or motivation); and, (6) gathering respondents' perceptions of an external environmental variable (e.g., supervisor's behavior, organizational climate, etc.). Categories 1 and 2 are unique from the other four because this data is verifiable through other means. Factual data can be cross-referenced with other records and experimental manipulations can incorporate checks to ensure the data is viable. Categories 3-6, however are subject to validity problems because they may not be verifiable through other means.

When using self-reports to measure these latter categories, respondents are asked to engage in higher-order thought processes. In addition to recall, these processes involve weighting, inference, prediction, interpretation, and evaluation. Respondents also are asked, in many cases, to think at a high level of abstraction. Podsakoff and Organ (1986) observed that by this time, the data obtained are several steps removed from the level of discrete stimuli and responses. These challenges can be overcome, however, provided that a measure of validity is ascertained.

Severe validity problems can ensue, however, when measures of two or more variables from categories 3-6 are collected from the same respondent and then attempts are made to interpret correlations among them (Podsakoff & Organ, 1986). When more than one measure is used to collect data from a single subject, common method variance

can occur (Podsakoff & Organ, 1986). When leadership ratings are collected from a source, and then effectiveness ratings are also collected from the same source, a defect in that source can compromise the results of both measures (Podsakoff & Organ, 1986). According to Podsakoff and Organ (1986), because rating source contamination generally occurs in the same fashion and in the same direction, correlations can appear more substantial than they really are.

Podsakoff and Organ (1986) explained several sub-forms of common method variance. When answering a series of items, respondents appear to maintain a consistent line of logic (or a line that appears consistent to the respondent). This *consistency motif* creates problems because people have developed personal theories on how behavior, personality, and environments inter-relate. In individual psychology, Adler referred to this interpretive dynamic as an individual's private logic (Ansbacher & Ansbacher, 1956). If respondents were reporting on discrete events then recall would be less susceptible to distortion, but the self-report data categories represented in 3-6 often call for respondents to provide summary judgments (Podsakoff & Organ, 1986). These judgments are the product of the individual's private logic. A respondent will interpret experience through this filter, generate associations, and then report accordingly. These associations can cause empirically distinct constructs to appear empirically related when this relationship actually is an artifact of the research design (Podsakoff & Organ, 1986).

A second form of common method variance results from the problem of *social desirability* (Podsakoff & Organ, 1986). Individuals might respond to items in socially favorable ways. This bias can shift the distribution of responses in an upward fashion. An

upward shift can attenuate correlations by virtue of range restriction, but additional problems can result as well. Some responses can be more ego-flattering than others. For example, if a hall director (HD) endorsed items indicative of high levels of job-related stress, then the HD may be more apt to respond to items implicating supervision deficiencies, incompetent resident advisors, or irrational policies or procedures than to acknowledge personal challenges such as poor planning or an inability to function in a team-oriented environment.

To control for common method variance limitations, Podsakoff and Organ (1986) recommended incorporating multiple assessment methods and multiple sources into research designs. If a multiple method, multiple source design is not possible, then researchers may use *post-hoc* statistical procedures to correct for the influence of covariance between measures. They provided a critical review of Harmon's one factor test (example provided in Greene & Organ, 1973), partial correlation procedure (e.g., Organ & Greene, 1981), elimination of social desirability (e.g., Crowne & Marlowe, 1964; Edwards, 1970), and scale item trimming (e.g., Birnbaum, Farh, & Wong, 1986). Researchers also may incorporate procedural interventions into their analysis. Podsakoff and Organ provided critical reviews of procedural interventions such as escalating the unit of analysis (e.g., Smith, Organ, & Near, 1983; Avolio, Yammarino, & Bass, 1991), separation of measurement, and scale reordering (e.g., Salancik & Pfeffer, 1977).

Summary for conditions of validity and limitations of self-reported data.

Self-report data collection is prevalent, cost-effective, and easy to administer. For these reasons alone, it is likely to continue to be a widely used method of data collection.

Certain precautions should be taken into account, however, when self-report methods are employed. The measurement conditions under which self-report data is collected can have a discernible impact on results (Mabe & West, 1982). Mabe and West (1982) cautioned that: (1) validity can be increased when respondents expect that a criterion variable will be measured in conjunction with their responses; (2) instructions should provide terminology that connotes relative, social comparison standards rather than absolute standards; (3) respondents' experience with self-evaluation can enhance or limit validity—opportunities to practice self-evaluation can enhance response validity; and, (4) participants should be informed that their responses will be anonymous. Further, Podsakoff and Organ (1986) identified six general categories under which self-report data is collected. The majority of these categories present limitations that should be addressed either through multi-trait, multi-method design or through post-hoc statistical procedures or through procedural interventions. Failure to address measurement conditions or neglecting to implement procedures controlling for statistical artifacts characteristic of self-report data will reduce the validity of self-report results.

Comparisons of Self-reported Data to Observer-reported Data

Multi-rater meta-analysis incorporating peer and supervisor ratings. Harris and Schaubroeck (1988) extended the research initiated by Mabe and West (1982). Whereas Mabe and West's meta-analytical study did not separate out self-peer and self-supervisor correlations or address peer-supervisor correlations, the current study sought to report a definitive estimate of rater agreement. Harris and Schaubroeck selected only studies that contained reliability estimates calculated according to accepted formulas so

that more accurate mean reliability estimates could be calculated across the studies. They also excluded studies that were conducted in laboratory environments. When multiple measures of performance were used, they averaged effect sizes (because most performance measures were not independent) and the mean effect was used in the meta-analysis. These procedural criteria yielded 36 independent self-supervisor correlations, 23 independent peer-supervisor correlations, and 11 independent self-peer correlations. Among the research questions investigated were the following: (a) Overall, what is the average correlation between self-supervisor, self-peer, and peer-supervisor ratings? (b) What is the average difference between supervisor, peer, and self-ratings? (c) Will the data support an egocentric bias, organizational-level, or observational opportunities explanation for rater disagreement?

The peer-supervisor correlation ($\rho = .62$) was notably higher than either the self-supervisor ($\rho = .35$) or the self-peer correlation ($\rho = .36$) (Harris & Schaubroeck, 1988). This finding indicated higher rater agreement among peer-supervisor observers than self-supervisor or self-peer groups. On average, self-ratings were found to be half a standard deviation higher than supervisor ratings and approximately one-quarter a standard deviation higher than peer ratings. Although two forms of egocentric bias were supported as an explanation for rating group discrepancies, evidence for organizational-level or observational opportunities was not found. The forms of egocentric bias that were supported were explained by attribution theory and through an explanation of moderated defensiveness. According to attribution theory, actors (self-raters) attribute good performance to their own behavior and poor performance to environmental factors (De

Vader, Bateson & Lord, 1986). On the other hand, observers (peers and supervisors) attribute good performance to environmental factors and poor performance to the actors' characteristics or behaviors. Since self-raters and observers (whether peer or supervisor) attribute performance to different sources, self-ratings should correlate poorly with observer ratings. Correcting for range restriction of upwardly biased self-ratings will have little effect on this outcome (they did not) and there should be substantial agreement between observer groups (there was). Egocentric bias also was supported through a theory of moderated defensiveness. Harris and Schaubroeck (1988) noted this theory postulated that defensiveness is moderated by other variables (Baird, 1977; Kay, Meyer, & French, 1965). Results from this meta-analysis suggested that work context could moderate defensiveness. The work contexts investigated were managerial/professional settings and blue-collar/service settings. Researchers found less bias evident in more concretely defined blue-collar/service settings than in the more abstract, less well-defined professional settings. Therefore, Harris and Schaubroeck (1988) proposed that egocentric bias is more likely to occur in more ambiguous jobs than in well-defined work settings. Because job type did not affect observer agreement, there was greater validity in observer ratings.

Multi-rater meta-analysis incorporating peer, superior, and supervisee ratings. Conway and Huffcutt (1997) examined the psychometric properties of subordinate, supervisor, peer, and self-ratings through the use of meta-analytic methodology. This study extended the previous work of Mabe and West (1982) and Harris and Schaubroeck (1988) by updating the studies under consideration as well as

including subordinate ratings within its analytical framework. Conway and Huffcutt sought to confirm whether multi-source ratings were valid performance measures. In particular, they studied how each rating source (subordinate, peer, supervisor, and self-ratings) contributed toward a better understanding of leader behavior. The study included 281 coefficients from 177 samples with a total sample size of 28,999. They compared managerial and non-managerial job types and cognitive as well as interpersonal dimensions. Several correlations illustrated a fallible relationship between self-ratings and leadership constructs.

Across all studies, the level of agreement between self and other ratings was minimal. The correlation between self-ratings and subordinate ratings across all jobs (managerial and non-managerial) was .14 (across 26 coefficients of study); the correlation between self-ratings and peer ratings across all jobs was .19 (across 17 coefficients); and the correlation between self-ratings and supervisor ratings was .22 (across 50 coefficients). Conway and Huffcutt speculated the low correlations could be due to egocentric biases (Harris & Schaubroeck, 1988), effects of self-esteem (e.g., persons with high self-esteem tend to overrate yet persons with lower self-esteem tend to underrate; Atwater & Yammarino, 1992), or because self-ratings measure something unique and valid outside the scope of other-reported data. Regardless of the cause, the implication remains the same: leaders were poor evaluators of their own leadership. Similar findings have been reported by Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski (2001), and Lombardo and Eichinger (2000).

Implications of directional ratings. To this point, the validity of self-ratings has been a key issue of consideration. On average, self-reports tend to over-estimate respondents' abilities (Mabe & West, 1982; Harris & Schaubrock, 1988; Conway & Huffcutt, 1997), this is not true for everyone (Atwater & Yammarino, 1992; Kaplan & Kaiser, 2006; Kruger & Dunning, 1999; Lombardo & Eichinger, 2000). Some individuals who exhibit strong self-awareness characteristics manage to rate themselves approximately the same as their observers (Atwater & Yammarino, 1992), but other self-raters underrate their performance in comparison to observers (Atwater & Yammarino, 1992; Kaplan & Kaiser, 2006; Kruger & Dunning, 1999). Given this differential, is the direction of rating discrepancies significant?

The self-ratings of the unskilled and unaware. Studies by Kruger and Dunning (1999) covered three cognitive domains (humor, logical reasoning, and grammar) and were assessed over four separate experimental conditions. The number of participants ranged from a low of 36 to a high of 140. All participants were college students at a large Northeastern university. Compared to others (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988; Mabe & West, 1982), Kruger and Dunning suggested a somewhat different reason for inflated self-ratings. They postulated that limited metacognitive ability (the ability to know how well one performs, when one is accurately judging performance, and when one is likely to be in error) contributes to overestimation of performance and ability. In other words, if one does not have the requisite competence to perform a task, then that person will also be a poor estimator of ability and performance related to that task—both in terms of self and others.

Four predictions were researched and all four were supported. First, a metacognitive deficit (low competence) was found to predict inflated self-assessments. Second, a deficit in competence hindered the ability to recognize competent performance in others; this deficit also decreased one's accuracy at predicting or estimating other peoples' ability. Third, those who were not competent were less able to gain insight into their true level of performance than were their more competent peers. Finally, those who were not competent could gain insight about their shortcomings, but such insight only came after having first made them more competent.

Across all four studies, the overall mean estimate of self-rated ability ranged from the 64th percentile to a high ranking of the 71st percentile. On average, participants considered themselves to possess above average ability—a somewhat illogical finding that reflected a bias toward inflated self-ratings. Most intriguing was not that bias was present, but rather the level of bias that was distributed in the bottom quartile as compared to that of the top quartile and the consequential effects.

Regardless of the cognitive domain or the study in question, students in the bottom performance quartile grossly overestimated their test performance. In the humor test, they estimated their performance to be in the 58th percentile, but their actual performance was in the 12th percentile—an overestimation of 42 points. In a logical reasoning test, participants in the bottom quartile estimated their performance to be in the 61st percentile but their actual performance was in the 12th percentile—a 45-point differential. Similar discrepancies were replicated in the other studies. Conversely, those in the top quartile consistently underestimated their percentile rankings.

To test predictions 2 and 3, Kruger and Dunning (1999) embedded a sub-study into the experimental design of the grammar assessment. They wanted to explore whether exposure to low-quality, average quality, and high-quality work would have any bearing on participants' self-ratings of ability and performance. To reiterate, bottom quartile participants grossly overestimated their percentile rank, yet top quartile participants underestimated their percentile rank both in terms of ability and performance. Four to six weeks after this initial study, participants scoring in the bottom and top quartiles were invited back to engage in a follow-up study. Samples of previously completed grammar tests were presented to participants (these tests represented quality levels characteristic of each standard of deviation of performance). Participants were informed of the varying quality levels and were asked grade each test by estimating the number of questions each test-taker had answered correctly. After this procedure, participants were shown their original test once again, asked to re-rate their grammar ability and re-rate their test performance relative to peers using the percentile ranking system. They also re-estimated the number of questions they answered correctly. Those who scored well on the test were significantly better at recognizing quality work; moreover, after being exposed to the varying quality levels of work, they significantly revised self-report estimates of ability and performance to more accurately reflect their actual standing within the sample. An entirely different result occurred for participants in the bottom quartile. They were significantly less able to discern competent work than their peers in the top quartile. Furthermore, bottom-quartile participants did not gain insight into their own ability after having been exposed to the more competent work of peers. Not only did bottom quartile

participants fail to significantly revise their self-evaluations of where they ranked within the sample (they were off the mark by more than 50 points), they slightly raised their self-estimates!

Kruger and Dunning (1999) proposed that two separate dynamics appeared to impact self-ratings of ability and revisions of self-ratings. For the top quartile, after having been exposed to representative work samples, they raised their self-estimates to a level more in line with their actual ability. The authors proposed that the *false consensus effect* explained this tendency (Ross, Green, & House, 1977). Absent data to the contrary, participants will assume that others performed at approximately the same level of competence. Exposure to the work sample rectified this mistaken assumption. For those in the bottom quartile a different explanation was provided. Their lack of competence prevented them from recognizing competence in others (support for prediction 2). Additionally, this incompetence hindered their ability to learn from other, more competent performers. Even after witnessing others who had performed considerably better, they still maintained the mistaken assumption that they also performed well (support for prediction 3).

Although leadership percentile rankings cannot be so finely determined as experimental testing procedures, self-rater inflation tendencies are similarly reflected in leadership studies (Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski, 2001; Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988). Furthermore, in leadership studies that have investigated the consequences of over-estimating one's ability, findings parallel the performance profiles of the lower and upper quartile participants in Kruger and

Dunning's studies. Over-raters have exhibited limited ability to learn from their experiences; in extreme cases, an inability to learn from mistakes can contribute to career path derailment (Lombardo & Eichinger, 2000). Under-raters and accurate estimators have demonstrated enhanced ability to learn from their experiences (Atwater & Yammarino, 1992).

Fortunately, Kruger and Dunning also tested whether or not skill development had a beneficial impact on metacognition. Subjects who were trained in a skill (in this case, a follow-up logical reasoning training session) demonstrated significantly better ability to evaluate others' work (after training, they actually did as well as top-quartile participants) and calibrate their own self-ratings. The lower quartile still overestimated their performance relative to peers, but the gap differential was considerably smaller. The training program improved low performer's competence and improved their ability to recognize competence in others and (presumably) learn from this exposure.

Kruger and Dunning discussed implications related to their findings. As long as low-performing subjects were unaware their performance was sub-standard, they assumed they performed at an above-average level. Furthermore, they were neither cognizant of others' skills nor able to derive insight into their own shortcomings without having received specific, direct training on what competent performance looked like. Competent performers carried a different burden. They assumed that because they performed well then others performed equally well; this assumption made them blind to their comparative ability.

Kruger and Dunning asserted that feedback—more specifically, deficit of feedback—contributed to those who are “unskilled and unaware” (p. 1121). The authors noted a rich body of research indicating that people rarely receive negative feedback about their skills and abilities in everyday life (Blumberg, 1972; Darley & Fazio, 1980; Goffman, 1955; Matlin & Stang, 1978; Tesser & Rosen, 1975). Also, they noted that some tasks and settings inhibit people from receiving self-correcting information that might reveal shortcomings or weaknesses. Leadership would fall into this category. Finally, even if people receive negative or constructive feedback, the value of this feedback is mitigated if they remain unaware of why the failure occurred. Kruger and Dunning (1999, p. 1131) observed:

[T]he problem with failure is that it is subject to more attributional ambiguity than success. For success to occur, many things must go right: The person must be skilled, apply effort, and perhaps be a bit lucky. For failure to occur, the lack of any one of these components is sufficient. Because of this, even if people receive feedback that points to a lack of a skill, they may attribute it to some other factor (Snyder, Higgins, & Stucky, 1983; Snyder, Shenkel, & Lowery, 1977).

Summary comparing self-reported data to observer-reported data. Multi-rater assessments provide a structured method for gaining multiple perspectives, for creating intentional feedback opportunities as recommended by Mabe & West (1982), and for doing so in a manner that is anonymous for those participating (Mabe & West, 1982; Upcraft & Pilato, 1982; Winston & Fitch, 1993). The validity of other-sourced data is greater than self-sourced data because individuals are prone to egocentric bias (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988), because self-ratings are

affected by the level of ambiguity associated with the rating context (Harris & Schaubroeck, 1988), and due to level of self-esteem (Conway & Huffcutt, 1997; Atwater & Yammarino, 1992). In other words, leaders are poor evaluators of their own leadership ability. In non-leadership studies, Kruger and Dunning (1999) demonstrated that a lack of competence was a predictor of inflated self-assessments. Unskilled performers were unable to perceive competence in others or gain insight into the level of their own performance. In contrast, skilled performers rated themselves lower than their ability should have suggested—this tendency was a result of the false consensus effect.

Collectively, these findings underscore problems associated with using self-only methods in leadership research. When accurate feedback is provided and training accompanies feedback, the unskilled and unaware demonstrated the potential to become competent and aware, and presumably more capable of learning from observation and from experience (Kruger & Dunning, 1999). Training programs designed around multi-rater protocols have demonstrated success for many decades (Eichinger & Lombardo, 2003; Leslie & Fleenor, 1998; McCauley & Van Velsor, 2004; Parry, 2005). Proper use of multi-rater assessment methods in university and collegiate settings can enhance the quality of research produced and when combined with effective feedback delivery, multi-rater assessment can also improve student outcomes (Atwater & Yammarino, 1992; Komives, 1991b; Posner, 2004; Posner & Brodsky, 1993).

Levels of Analysis

Pascarella and Terenzini (2005) reported that since 1990, the majority of leadership research has addressed institutional factors that influence leadership

development. These studies have examined: developmental outcomes of college students' involvement (Astin, 1993; Sax & Astin, 1998; Cress, Astin, Zimmerman-Oster, Burkhardt, 2001); impacts and insights from U.S. college and university leadership development programs (Zimmerman-Oster & Burkhardt, 1999); a grounded theory of high quality leadership programs (Eich, 2007); college student perceptions of leadership and contributing environmental factors (Multi-institutional Study of Leadership [MSL], 2006; Shertzer & Schue, 2004; Thompson, 2006); institutional spending patterns and leadership outcomes (Smart, Ethington, Riggs, & Thompson, 2002); and, validity of a college-focused leadership model that attempted to integrate individual level, group-level, and society-level values into a common framework and assessment instrument (Dugan, 2006a; Dugan 2006b; Higher Education Research Institute [HERI], 1996; MSL, 2006; Tyree, 1998). This list is not exhaustive, but it represents a significant balance of leadership publications within contemporary student leadership research. There is a need to supplement the institutional focus of these studies with additional research that focuses on leader-member behaviors to better inform practices of leadership development in institutions of higher education.

The ecological fallacy. Researchers can draw erroneous conclusions, termed ecological fallacies, when they use aggregated leadership data to make inferences about individual level effects (Antonakis, Schriesheim, Donovan, Gopalkrishna-Pillai, Pellegrini, & Rossome, 2004). Focusing a majority of research on institutional characteristics while neglecting to balance it with studies on interpersonal behavior might be the equivalent of studying an ecosystem without analyzing relationships among the

species therein. A macro analysis is valuable, but it should be informed with micro-level hypotheses and research methods. Depending the level of analysis that has been used, research conclusions may differ because “the correlates or causes of individual performance may be very different from the correlates or causes of group or organizational performance” (Antonakis, Schriesheim, et al., 2004, p. 63). Studying institutional characteristics increases awareness of institutional factors, but this institutional agenda is not a sufficient substitute for interpersonal leadership research. The institutional research agenda and the interpersonal research agendas should complement one another, for neither suitably replaces the other.

Pascarella and Terenzini (2005) observed that the majority of student leadership studies were self-report designs or studies using pre-existing institutional data, such as Smart, et al., (2002). The validity of these kinds of studies is subject to question. (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988; Mabe & West, 1982; Podsakoff & Organ, 1986). Because leadership is a social behavior, not an individual one (Denison, Hooijberg, & Quinn, 1998), self-report research design is limited in its ability to measure interpersonal aspects of leadership because self-report-only designs are blind to other peoples’ perceptions of the leader—and observers are the most valid predictors of leadership effectiveness (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988; Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski, 2001).

Multi-rater methods allow for segmented analysis of varying levels of leadership interactions. Insights that would not be accessible at the overall-level of analysis can be gleaned using a multi-rater, multi-level analytical design. Atwater and Yammarino’s

(1992) multi-rater study on U. S. Navy cadets demonstrated how multi-rater assessment methodology, coupled with appropriate analysis of different rater levels, yielded different results than would have been obtained if analysis were conducted at the overall level.

Self-awareness and multi-rater methodology. Many studies of self-reporting-on-self assessment address the degree to which self-reported data provides meaningful information about the self-rater as a source of information, e.g., Conway and Huffcutt (1997). Atwater and Yammarino took a different approach. They explored how self-ratings, compared with observer ratings, were indicative of the focal leader's self-awareness. Citing Wicklund's self-awareness theory (1975; 1978; 1979), they posited that self-awareness emanates from one's ability to self-observe. Those who are self-aware will compare what they know about themselves to new information or external benchmarks. Highly self-aware individuals integrate information from these assessments more effectively than those lower in self-awareness. As a result, a leader's self-evaluation and behavior will mature as a result of iterative analysis and integration. As individuals become more self-aware, they become more cognizant of how they are perceived by others. Consequently, individuals with high self-awareness should exhibit greater self-other rater agreement than those with lower self-awareness. Atwater and Yammarino (1992) considered the following research questions: (1) To what extent do individuals inflate self-ratings of leadership? (2) What contributes to inflated self-ratings of leadership? (3) How does self-awareness (defined as agreement between self- and other ratings) affect the degree of relationship between leader behavior and performance outcomes?

Atwater and Yammarino (1992) studied a sample of upperclassmen at the U. S. Naval Academy (19-22 years of age). The sample included 91 student leaders rated by 1,145 freshman (subordinate raters) and 11 military officers (superior raters) responsible for the student leaders. It should be noted the sample was heavily biased toward men; only 8 of the 91 student leaders were female. Atwater and Yammarino used the self, observer, and superior forms from the Multifactor Leadership Questionnaire (MLQ) (Bass, 1985; Bass & Avolio, 1990) to assess student leadership behavior. Unlike many studies with the MLQ, they used only the items from the four transformational leadership scales. In addition, because of high intercorrelations among subscales in their sample (.71 to .88), Atwater and Yammarino treated the 24 transformational items as if they were one scale rather than as the separate subscales typically identified in MLQ-related research. Observer scores for each category of raters (subordinate or supervisor) were averaged into a single rating. The appropriateness of this procedure was evaluated through a one-way ANOVA (see Sheridan & Vredenburgh, 1978) and Bartlett's M-test examined the homogeneity of within-leader variance. They measured *student ability* through SAT scores, student application recommendation forms, as well as scores on the engineering and science interest scale. *Student experience* was measured through athletic participation, leadership positions held, and conduct records.

Since self-ratings tend to be inflated when compared to observer ratings (Harris & Schaubroeck, 1988), it is difficult to directly compare mean scores between self-rater groups and observer groups. To overcome this limitation, Atwater and Yammarino (1992) categorized self-raters into one of three agreement groups according to the

magnitude of difference between self-scores and observer scores. The researchers calculated a distribution of difference scores for each self-other comparison. Those whose self-scores deviated less than one-half a standard deviation from the mean difference for that comparison were identified as *in-agreement*. Those whose difference scores were one-half standard deviation or more above or below the mean difference were respectively identified as *over-estimators* or *under-estimators*.

Atwater and Yammarino's (1992) findings supported previous literature that found some individuals inflate their self-ratings of leadership (Harris & Schaubroeck, 1988). This inflation emanated from two sources. First, self-raters overestimated their leadership; second, higher self-scores were correlated with lower observer scores, contributing to greater disparity between self versus other scores. Inflation was not merely a product of the leader, but also resulted from the observers' perceptions of the leader (Atwater & Yammarino, 1992). This study provided evidence that self-awareness can moderate the relationship between leader behavior and how it corresponds with predictor variables and performance (Atwater & Yammarino, 1992).

Looking at broad trends within the data for each agreement category, every significant correlation of ability and leadership was negative except one. The over-estimator group, however, was most impacted by negative correlations. Atwater and Yammarino (1992) speculated that the self-aware group and under-estimator group may have learned lessons from their experiences, modified their behavior, and improved their leadership scores thereby lessening the impact. Such learning was evidenced by the presence of five significant, positive correlations between leadership scores and

experience variables. The opposite was true for the over-estimator group. Negative correlations of ability were not offset by positive correlations between leadership and experience. To the contrary, the correlation between leadership ability and leadership positions held provided the highest negative correlation within the analysis (-.58).

Although the over-estimator group possessed high ability metrics, these ability scores did not translate well to leader-member relationships. Experience did not assuage these limitations for over-estimators; in fact, leadership experience exacerbated them.

Atwater and Yammarino (1992) referenced Ashford's (1989) observation that early, formative leadership experiences tend to infuse leaders with beliefs about themselves. Leaders filter subsequent experiences and information according to these beliefs. Apparently, leaders who possess high ability scores and who over-estimate themselves may create a self-portrait that is out of step with others' perceptions of their leadership. They risk interpreting their leadership experiences through a more forgiving filter than observers do, attributing successes to their ability and attributing failure to others' fallibility (DeVader, Bateson, & Lord, 1986; Harris & Schaubroeck, 1988; Martinko, Harvey, & Douglas, 2007). In addition to these interpretations, the data provide leadership-specific support for findings in Kruger and Dunning's (1999) study. Limited competence leads to inflated self-assessment, predicts poor self-calibration, and these individuals gained limited insight from observation or experience. Limited, that is, until awareness was stimulated and competence could be achieved. Although these unskilled and unaware leaders rated themselves as good leaders, those working with them

believed otherwise. As Kruger and Dunning (1999) noted, these leaders will not learn to be more effective without constructive feedback.

Findings in this study (Atwater & Yammarino, 1992) have serious implications for student leadership development. When high self-esteem is not accompanied by high self-awareness, leaders exhibit a tendency to overrate their leadership abilities. Therefore, leaders can form misguided impressions of their leadership ability, filter their experiences through faulty lenses, and continue overlooking the learning opportunities presented through leadership experience. Alternative sources of feedback, such as multi-rater instrumentation, can promote constructive change in self-perspectives (Parry & Sinha, 2005). As Kruger and Dunning (1999) observed, such feedback is an essential component of promoting competence. Participants in this study were provided feedback on their results and their response was overwhelmingly positive (Atwater & Yammarino, 1992). Because multi-rater feedback is anonymous, systematic, and structured, it is often met with greater receptiveness than other types of feedback (e.g., performance reviews) (Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski, 2001; Harris & Schaubroeck, 1988).

A multi-rater design incorporating levels of analysis was necessary to yield the findings of this study. A self-only research design would not have been sensitive to the leader-member dynamics revealed in this study. In fact, the least effective leaders would have provided the highest leadership ratings by virtue of an inflated sense of their leadership skills—even though criterion variables would have seemed to contradict this finding. Observer responses were necessary to put self-ratings into context. Furthermore,

had Atwater and Yammarino (1992) chosen to review overall data rather than subdividing self-raters into agreement categories, their findings would have appeared dramatically different. Correlations extracted at the overall level of analysis did not parallel results from any of the agreement categories (Atwater & Yammarino, 1992) In subordinate ratings, for example, leadership positions were significantly and positively correlated with leadership scores for the in-agreement group. In the over-estimator and under-estimator groups, however, leadership positions were negatively and insignificantly correlated with leadership scores. Consequently, in the overall analysis, leadership position did not appear to be significant because the negative, insignificant correlations overpowered the positive, significant correlation for the in-agreement group. Leadership researchers who fail to take levels of analysis into account risk “building theoretical skyscrapers on foundations of empirical jello” (cf. Schriesheim, Castro, Zhou, & Yammarino, 2001, p. 516).

Summary of levels of analysis. Pascarella and Terenzini (2005) indicated that college has a beneficial impact on students’ leadership skill development, but specifically how and where students develop these skills remains unclear. Research has yet to cull specific, significant entities of impact. Student leadership literature has focused heavily on environmental influences and has attempted to measure leadership growth through self-reported outcomes based on students’ thoughts or feelings about leadership rather than on observable behaviors. In other words, this research has focused on how the environment affected student leaders’ thoughts and feelings, not on how student leaders behaved and thus impacted their environment. Consequently, this research has provided

little insight into where and how students improve their leadership skills. The strong reliance on self-report data not only limits the validity of data, but also restricts the ability to analyze data from multiple levels. Multi-rater, multi-level studies permit greater insight into developmental needs and characteristics of select groups—needs and characteristics that would go unnoticed at higher levels of analysis. Atwater and Yammarino's (1992) study demonstrated how levels of analysis can uncover significant relationships that remained hidden when analyzed at an aggregate level.

Development of Leadership in Housing and Residence Life Contexts

Self-reported measures need to be complemented by alternative sources of data (Podsakoff & Organ, 1986). Levels of analysis should be incorporated into research designs because findings that are significant at one level may not be at another (Yammarino & Atwater, 1992; Yammarino, Dionne, Chu, & Danserau, 2005). Research designs should use measures that are appropriately suited for the task and utilize analytical procedures appropriate for that level (Kroeck, Lowe, & Brown, 2004; Yammarino, Dionne, Chu, & Danserau, 2005). Before accepting Pascarella and Terenzini's (2005) assertion that skill development results from cumulative effects rather than being a product of context-specific experiences, researchers should design studies that are behaviorally focused and that analyze student leadership and outcomes at multiple levels. Therefore two questions should be answered: First, what processes, procedures, and instruments would be most effective for studying a specific context?; Second, what student leadership context would be appropriate for further study?

At the present time, these two questions are interdependent. Regardless of the context being studied, very few leadership instruments have been designed and validated for college campuses (Posner, 2004; Schwartz & Gimbel, 2000). A multi-rater instrument would be preferable because it would maximize validity and allow for multi-level analysis, but the Student Leadership Practices Inventory (SLPI) is the only student-focused, multi-rater instrument designed for and validated on a student population (Leadership Challenge, 2008). Although using the SLPI could suffice, expanding the range of multi-rater instruments could be a valuable contribution to the student leadership literature. The present study seeks to broaden the range of multi-rater instruments by providing an instrument that measures the versatility of young leaders.

To validate a new behavior-based, multi-rater leadership instrument on a college population, context becomes an important question. Many college leadership domains are loosely structured (e.g., student organizations, project groups, service teams) and leadership might be the product of emergent processes rather than from a formally assigned leadership role. Such fluidity may or may not be a present problem when using established instruments, but a loosely structured rating context could pose problems when attempting to validate a new instrument (e.g., ambiguity of work context was a mediating factor in Harris and Schaubroeck, 1988). Ideally, observers and leaders need to have a clear understanding of their respective roles, and have prior experience working with them in a formally defined leader-member relationship.

Student Leadership Context: Housing and Residence Life

To identify an environment well suited for a leadership validation study, Komives and colleagues provided helpful guidelines in their research on student leadership identity formation (Komives, Longerbeam, Owen, Mainella, & Osteen 2006; Komives, Owen, Longerbeam, Mainella, & Osteen 2005). Their six-stage model provided methodological insight into conditions and experiences that promote student leadership growth.

According to this research, students who participated in leadership activities that involved adults, peers, meaningful roles, and reflective learning practices experienced greater opportunity for personal growth. Furthermore, college and university personnel who provide such opportunities for students can improve growth and development outcomes through intentional assessment, challenge, and support of participants (Komives, Longerbeam, et al., 2006; McCauley & Van Velsor, 2004).

Paraprofessionals in housing and residence-life. Housing and residence life is one such context where students can learn and grow as leaders (Winston & Fitch, 1993). Students working in housing and residence life have opportunities for adult engagement through relationships and interactions with professional housing and residence-life administrators. Paraprofessional roles, in particular, are well suited for student leadership development. Winston and Fitch (1993) described housing and residence life paraprofessionals as students who were hired, trained, and supervised to assume responsibilities and perform tasks that promote the development of peers, create and maintain stimulating living environments that support residents' personal and educational development, and take appropriate measures to ensure safe, clean, healthy,

psychologically safe, and esthetically pleasing residential accommodations. Winston and Fitch (1993, p. 317) further noted:

[A] crucial dimension of this definition is the essential roles that selection, training, and supervision play: in the absence of carefully conceptualized, expertly implemented, and systematically pursued training and supervision, potential paraprofessionals simply become at best well-intentioned amateurs whose successes are attributable mainly to forceful personalities and/or fortuitous circumstances.

Both undergraduate and graduate students can occupy paraprofessional roles.

The Resident Advisor role. The most prominent undergraduate paraprofessional role is the Resident Advisor (RA) position. Blimling (1998) identified five key roles that RAs negotiate: student, role model, counselor, teacher, and administrator. RAs are students first, but often struggle with neglecting their grades because of the all-consuming nature of the position (Blimling, 1998; Paladino, Murray, Newgent & Gohn, 2005). RAs are role models and the behaviors they model, whether in or outside of their residence hall or not, are highly influential. They must possess excellent self-monitoring skills; consequently, RAs are expected to maintain a sense of decorum and composure regardless of where they are on campus (Blimling, 1998; Upcraft & Pilato, 1982). In the counseling role, Blimling (1998) stated that RAs are front-line interventionists for students in need of help. RAs serve as teachers, through direct programming and group facilitation, or how they display and convey their values on the residence hall (Blimling, 1998; Upcraft & Pilato, 1982; Winston & Fitch, 1993).

RAs also administrate. They keep organized records, ensure facilities are appropriately maintained, and that students receive the services they need from housing

and residence life (Blimling, 1998). Even though the RA position is well suited for the development of student leaders, the quality of an RA's experience depends upon the support and training they receive from superiors (Murray, Snyder, & Midkiff, 1999; Paladino, Murray, Newgent & Gohn, 2005; Upcraft & Pilato, 1982; Winston & Buckner, 1984). If RAs are not effectively trained and supported, the responsibilities they manage can overwhelm them (Ellevan Allen, & Wircenski, 2001; Deluga & Winters, 1990; Paladino, Murray, et al., 2005).

The Hall Director role. Both professional-level student affairs personnel and graduate students serve as supervisors for RAs. Many terms have been used to describe this supervisory role (e.g., Community Advisor, Community Director, and Hall Coordinator); in this study they will be identified as Hall Directors (HD). HDs are responsible for oversight and management of residence halls (Winston & Fitch, 1993). HD responsibilities include management of a residential unit with sizes ranging from 100-500 beds, structuring the educational and psychological environment of the residence hall, maintaining discipline, facilitating administrative functions, counseling troubled residents, encouraging social and co-curricular programming, and providing training and supervision for RAs (Winston & Fitch, 1993).

Winston and Fitch (1993) noted that Master's-level HDs may find their position somewhat ambiguous, "they are more mature and have greater responsibilities than RAs, but they lack the status, authority, experience, professional academic preparation, and income of well-qualified housing professionals ... Consequently, they frequently are neither fish nor fowl, neither paraprofessional nor professional, but some amalgam of the

two” (pp. 317-318). Master’s level HDs require as much training and supervision as RAs, although the supervisory aspect of HD responsibilities requires different approaches toward developing effective HD skills (Winston & Fitch, 1993).

According to a study by Winston and Fitch (1993) many Master’s level HDs are enrolled in student affairs preparatory programs. These HDs generally valued their position more than other HDs because their position provided a professional development context where classroom learning could be applied and students could gain relevant career experience (Winston & Fitch). The most difficult challenges that HDs coped with were time management (managing job, academics, and personal life), maintaining discipline and enforcing rules, managing challenges associated with living on a residence hall (lack of privacy, noise, etc.), attending meetings, and keeping up with paperwork (Winston & Fitch, 1993).

The literature specific to HD training and development is sparse. A number of resources have been written on the supervision, training and development of RAs (Upcraft & Pilato, 1982; Winston & Fitch, 1993; Blimling, 1998) and of student affairs professionals in general (Dirkx, Gilley, & Gilley, 2004; Holmes, 1998; Janosik, Creamer, Hirt, Winston, Saunders, & Cooper, 2003; Roberts, 2007; Winston & Creamer, 1997; 1998). Because the Master’s level HD is “neither fish nor fowl” (Winston & Fitch, 1993, p. 318), the training and development needs of non-professional HDs must be deduced from literature on RAs and student affairs professionals. Therefore, the literature on leadership and supervision in residence halls has been divided into a section that addresses RA training and development issues and a section that addresses professional

student affairs training and development issues. A summary of these two bodies of literature will identify common themes that are relevant to both. These common themes will then be used as a basis for HD training and development needs.

Leadership Styles and Leadership Development in Residence Hall Settings

The housing and residence life leadership literature draws upon behavioral, situational (Upcraft & Pilato, 1982; Winston & Fitch, 1993; Winston, Ullom, & Werring, 1984), and transformational (Brodsky & Posner, 1993; Komives, 1991a; 1991b; 1991c; Posner, 2004) leadership theories. Although much of this literature refers specifically to the leadership roles of RAs, many of these same interpersonal leadership principles are applicable to other residence life paraprofessionals such as HDs (Upcraft & Pilato, 1982; Winston & Fitch, 1993).

General Observations About Paraprofessional Residence Hall Leadership and Governance

RAs are only as good as their training and supervision (Upcraft & Pilato, 1982; Winston & Fitch, 1993). The HD delivers a critical service toward providing effective residence hall leadership (Upcraft & Pilato, 1982; Winston, Ullom, & Werring, 1984; Winston & Fitch, 1993). Evaluation is key to effective supervision. Participants should believe that the evaluation process open, fair, and valid (Upcraft & Pilato, 1982; Winston & Fitch, 1993). Effective evaluation systems provide a clear statement of job functions and expectations, identify the sources of information used for evaluation purposes, conduct an informational conference that reviews the job expectations and evaluation processes, and evaluate the leader based on the stated criteria (Upcraft & Pilato, 1982).

When evaluating performance, information should be collected from multiple sources—the constituents involved, the leader, and from the leader’s supervisor (Upcraft & Pilato, 1982; Winston & Fitch, 1993).

Generally speaking, high authoritarian leadership styles do not work well in residence hall environments; conversely, students who have difficulty exercising authority and accepting responsibility will have difficulty as well (Upcraft & Pilato, 1982; Winston, Ullom, & Werring, 1984). The dominant leadership style exercised by a residence hall leader will influence the social climate of the living unit and will establish the style of interaction leaders have with their constituents (Winston & Fitch, 1993).

Training and development efforts should use interventions designed to help residence hall leaders become more flexible in their leadership styles (Upcraft & Pilato, 1982; Winston & Fitch, 1993). The most effective leaders can accurately assess the situation, and then utilize the leadership style that is most likely to fit that situation (Upcraft & Pilato, 1982). Citing Deluga (1989), Winston and Fitch reported findings that indicated the residence hall leader’s ability to “generate and maintain a favorable group atmosphere, regardless of task-relationship leadership orientation, may prove to be a primary factor affecting ... influence patters” (p. 326).

Interpersonal Characteristics and Skills of Paraprofessionals

Leaders must be attentive to nonverbal behavior, actively listen, and choose appropriate methods for disclosure of thoughts and feelings (Upcraft & Pilato, 1982; Winston & Fitch, 1993). At times, leaders will need to initiate behavior, summarize events that have occurred, and confront incongruence when people say one thing but are

doing another. These skills are equally necessary for one-to-one and group leadership situations (Upcraft & Pilato, 1982). Some individuals excel in one-to-one relationships, but struggle when faced with group-level interactions (Upcraft & Pilato, 1982). Upcraft and Pilato (1982) identified several skills that were absolutely key for effective residence-hall leadership: Learning when to act in a group intervention and when to hold back; learning how to make a group aware of its own dynamics when doing so will help the group; the ability to recommend strategies to help groups move forward to accomplish goals; effective conflict mediation; and helping groups develop skills, such as feedback delivery, that enhances communication among group members.

A recent study from Jager and Caison (2006) examined the relationship between emotional intelligence and outstanding RA performance. Adaptability, problem solving, and flexibility were three of the most significant predictors of RA excellence. They posited that learning which competencies are characteristic of outstanding RAs can help housing and residence life administrators with candidate selection and training (Jager & Caison, 2006). This observation parallels the findings presented in professional leadership literature, e.g., Lombardo & Eichinger (2000).

Paraprofessional Role Clarity

Residence hall leaders need to have accurate knowledge of both the potential and limitations of their positions when leading and working with groups (Upcraft & Pilato, 1982). Both RAs and HDs have positional authority in their roles. Upcraft and Pilato (1982) referenced French and Raven's (1960) five sources of power leaders use to influence others: referent power (influence through modeling appropriate behavior),

legitimate power (positional influence), expert power (content knowledge and interpersonal influence skills), reward power (power of positive reinforcement), and coercive power (ability to influence through punishment or discipline). Power and influence needs to be used with discretion; knowing the scope of one's role and responsibilities helps inform a leader's judgment (Upcraft & Pilato, 1982).

Winston and Fitch (1993) stated that housing programs should create clear and unambiguous statements of position responsibilities and expectations. One of the most frequent challenges RAs encounter is deciding whether or not to act in a specific situation, and if they do need to act, what course of action should they take. According to Perry's (1970) developmental scheme, traditionally aged RAs have difficulty distinguishing between competing points-of-view on an issue (Winston & Fitch, 1993). Clarified roles and responsibilities help students navigate gray areas and the competing opinions RAs typically face (Winston & Fitch, 1993). Winston and Fitch (1993) noted that structure can provide RAs confidence in their positional role and responsibilities.

Paraprofessional Supervisory Styles

Insecurity can complicate the HD/RA supervisory relationship. Winston and Fitch (1993) observed that many RAs are supervised by entry-level professionals lacking basic supervision skills and have little confidence in their abilities or knowledge of their role or authority. In professional leadership literature, supervisors who have insecurities typically respond by overcompensation or by avoidance (Kaiser & Kaplan, 2006). Adler observed this same dynamic in early psychological literature. He identified it as the inferiority complex (Ansbacher & Ansbacher, 1956). In other words, insecurity prompts

people to either under-use a behavior, avoiding it because they are discouraged about their ability to succeed, or to overuse it, and thus compensate for feelings of insecurity or inferiority.

As Adler noticed in the general population (Ansbacher & Ansbacher, 1956) and Kaplan and Kaiser (2006) observed in executive leaders, supervision of RAs tends to fluctuate between extremes (Winston & Fitch, 1993). Some HDs exercise little to no control and have limited knowledge of the RAs activity on their unit. On the other hand, some supervisors are stifling in their control, giving RAs little to no latitude in their decision-making (Winston & Fitch, 1993). These are classic examples of an underdo / overdo reaction to leadership role insecurity (Kaplan & Kaiser, 2006). Distinct supervision styles have been identified and the benefits and tradeoffs of each have been documented.

The Authoritarian Supervisor. Many problems with RA supervision rest with the Hall Directors themselves. Supervisors who are overly authoritarian or paternalistic—not trusting RAs to work independently—will monitor RAs very closely (Upcraft & Pilato, 1982; Winston & Fitch, 1993; Winston, Ullom, & Werring, 1984). They tell RAs what to do, when to do it, and how. If RAs do not comply, then the authoritarian HD is likely to punish as a means to restore compliance (Upcraft & Pilato, 1982). In response to this supervisory style, RAs will conceal what is really going on because the HD's recommendations may not always work on a specific floor or for unique situations (Upcraft & Pilato, 1982). An “us” versus “them” mentality may be fostered, where the

RAs and students conspire against the residence hall administration (Upcraft & Pilato, 1982; Winston & Fitch, 1993).

The Laissez-faire Supervisor. The laissez-faire approach is a more frequently used style than the authoritarian approach (Upcraft & Pilato, 1982). Laissez-faire is the opposite of authoritarian—hands-off until absolutely necessary (Upcraft & Pilato, 1982; Winston, Ullom, & Werring, 1984). Training is infrequent and uncoordinated; the operating assumption is that only RAs can really determine how to work effectively on their floors because they live there while supervisors do not. Supervisors may serve as a sounding board, but are not partners in decision-making processes (Winston & Fitch, 1993). If problems do not arise, then the RAs and supervisors have very little contact. Consequently, RAs might do very little and get away with it; furthermore, because of the discontinuity of this approach there is little uniformity in the way different RAs operate or perceive their roles (Upcraft & Pilato, 1982).

The RA's RA. A third leadership style is exemplified when a supervisor tries to become the RA's RA. Winston, Ullom, and Werring (1984) termed this style maternalistic supervision. The supervisor establishes a warm, personal friendship with the RA and serves as a counselor, advisor, and friend. Most interactions focus on the supervisor's relationship with the RA and the RA's personal growth and development. The rationale behind this approach is that when supervisors model this approach, an RA will duplicate this relationship with residents. The problem is that the supervisor is focused on the best interests of the RA, not on the best interests of the student residents. Upcraft and Pilato (1982) believe this approach does not work. Winston and Fitch (1993)

believed that this approach placed too much emphasis on the RA's individual needs and not enough emphasis on housing and residence life objectives.

The Synergistic Supervisor. Winston, Ullom, and Werring (1984) indicated that effective residence hall supervision was a dynamic balance of leadership qualities. Supervisors appreciate the personal challenges and triumphs of an RA but still maintain focus on the organizational objectives of the residence hall. RAs are encouraged to share problems and participate in the troubleshooting process; but when necessary, the HD will provide directive influence ensure the institution's interests remain in focus (Winston, Ullom, & Werring, 1984).

Winston, Ullom, and Werring (1984) extended Upcraft and Pilato's (1982) approach to supervision styles by adding a fourth supervision style: the synergistic style. The synergistic supervisor utilizes a collaborative approach, where the supervisor and the RA combine efforts in such a way that the sum of both is greater than the individual parts of either. The HD and RA strive to accomplish organizational goals as well as further the personal development of the RA. To prevent escalation of small problems into crisis-level issues, HDs encourage RAs to speak openly about events occurring on their living units. Problems that an RA brings to the HD are perceived as situations that require joint effort, not as failures on the part of the RA. This approach allows them to jointly consider strategies for individuals on the residential unit. The supervisor can adopt a consultative role, broadening the range of issues an RA might consider. When properly executed, the synergistic relationship generates co-constructed solutions. These solutions help address pending issues, but they also provide RAs opportunities for growth. RAs experience the

benefits of a measured approach toward leadership—not too authoritative nor too laissez-faire; the supervisor models how to balance between concern for the individual and maintaining respect for institutional objectives (Winston, Ullom, & Werring, 1984). In professional business literature, Martin (2007) addressed the importance of the “opposable mind” (p. 62). Successful leaders integrate ideas and possibilities that are seemingly at odds with one another (Martin, 2007). The concept of integrative thinking is not valuable, however, unless it is translated into meaningful behavior. Denison, Hoojberg, and Quinn (1995) noted, “cognitive complexity ... may well be a *necessary* condition for the effective practice of leadership. Behavioral complexity, however, must certainly be the *sufficient* condition. Leadership must inevitably be performed through action, not cognition” (p. 524). Kaplan and Kaiser’s (2006) LVI addresses integrative thinking on a practical front. It measures versatility through observable behaviors and the instrument reveals lopsidedness or versatility in a leader’s behavior. The synergistic supervisor, as described by Winston and colleagues, conceptualizes the cognitive complexity needed for effective supervision (Winston & Creamer, 1997; Winston & Fitch, 1993; Winston, Ullom, & Werring, 1984). It does not, however, provide a practical method for measures the behavioral correlates of synergistic supervision. It describes the behaviors, but not how to balance them or how to measure that balance.

Development and Supervision of Student Affairs Professionals

In terms of interpersonal behavior, a significant portion of the professional development and supervision literature mirrors the paraprofessional literature. Lovell and Kosteen (2000) analyzed 30 years of empirical research in student affairs professional

development. Of the 23 empirical studies reviewed, 78% addressed issues of human facilitation. Counseling and staff supervision skills were considered critical for success in the student development profession (Lovell & Kosteen, 2000). Specifically, Lovell and Kosteen (2000) referenced studies by Tillotson (1995) who emphasized the importance of interpersonal relationship skills, organization skills, communication skills, and the directive skills necessary for working with others. Some studies referenced leadership in more general terms, identifying leadership ability as one of the most important aptitudes for student affairs professionals (Kane, 1982; Fey, 1991, Gordon Strode, & Mann, 1993; Roe, 1981). Schreiber, Dunkel & Jahr (1994) recommended the use of competency models to guide student affairs professional development. Leadership, communication skills, and interpersonal processes were consistent themes in the proposed models (ACUHO-I, 2008; Schreiber, Dunkel & Jahr, 1994; Winston & Creamer, 1997). Developmental programs based on competency assessment, were recommended for the development of talent within the profession.

Problematic Approaches to Supervision

In paraprofessional literature, authoritarian leadership styles were identified as overly intrusive and disruptive (Upcraft & Pilato, 1982; Winston, Ullum, & Werring, 1984). This challenge was not limited to paraprofessionals. Professional supervisors may also struggle to balance between directive action and providing subordinates enough leeway to challenge themselves, make mistakes, and grow from experience (Arminio & Creamer, 2001; Winston & Creamer, 1997; Winston & Hirt, 2003). How supervisors choose to use their power has a direct impact on subordinates (Winston & Hirt, 2003).

Winston and Hirt (2003) referenced power-based insights from Yukl (1998); exercising power can produce commitment, compliance, or resistance. Commitment is produced when power is appropriately applied and staff is supportive of the objective. Compliance is produced when staff members are not in agreement with the task, but are willing to invest minimal effort to comply with supervisor wishes. Compliance affects behavior, but not attitudes (Winston & Hirt, 2003). Therefore compliance is less influential than commitment. Resistance is produced when staff members push back against a requests or proposals and actively avoid trying to carry it out. Winston and Hirt (2003) charted effective and ineffective uses of power. Through this chart, the authors articulated the consequences of overusing forceful behavior. Power can be viewed as a nonlinear influence—appropriate power can maximize influence and commitment, especially when applied at the right time and behind the right ideas. When overused, power can provoke compliant or resistant behaviors—neither of which is as effective as commitment. These observations are consistent with findings from contingency theory (Vroom & Yetton, 2003; House, 1971; House & Mitchell, 1974) and align with Kaplan and Kaiser’s (2006) versatile leader model. When using forceful behaviors, supervisors need to learn how to use power at the right time and toward the appropriate ends if they want to be as influential and effective as possible.

Another other notable problem in professional supervision was the absence of supervision (Winston & Creamer, 1997; 1998). Because of the hectic pace of student-affairs, it becomes easy for supervisors to overlook holding regular supervisory meetings. When problems escalate to the point they must be addressed, meetings will ultimately be

held. As noted in paraprofessional literature, however, this approach undermines the potential benefits of supervision. Subordinates view this type of supervision as remedial or punitive (Winston & Creamer, 1997). These supervisory interactions do not foster a sense of interdependent personal development and accomplishment (Winston & Creamer, 1997; 1998).

Systematic Development and Supervision of Student Affairs Professionals

Professional supervision in student affairs is focused on the development of talent so that staff members are better equipped to achieve their organizational mission (Winston & Creamer, 1997; 1998; Dalton, 1988; 1996). Schrieber, Dunkel, & Jahr (1994) defined systematic development as “involvement in activities that are intended to enhance professional effectiveness, and are chosen as a result of a decision-making process based on assessment of skills and designed goals while targeting skill development” (p. 26). Holmes (1998) addressed this same concept in student affairs, but under the broader terminology of human resource development. Supervision and professional development are considered to be integral processes (Arminio & Winston, 2001; Holmes, 1998; Winston & Creamer, 1997; 1998). Effective supervision recognizes and addresses interpersonal needs of the supervisee, but also considers these needs within the performance context required to fulfill institutional responsibilities (Armino & Winston, 2001; Winston & Creamer, 1997; 1998; Winston & Hirt, 2003).

A variety of approaches have been recommended for development. Group-based approaches include in-service training, professional conferences, and graduate academic programs and individual approaches include discussions with colleagues, mentoring, and

individually focused programming (Roberts, 2007). Specialized workshops have demonstrated success, as exemplified by the eighteen-year history of the National Housing Training Institute (NHTI) (NHTI, 2008). The NHTI program has utilized Dunkel and Schrieber's (1990) competency model (updated by Porter [2005]) to help participants learn more about each competency, assess their abilities, and design a personal development curriculum (NHTI, 2008). In a survey of student affairs professionals, Roberts (2007) respondents indicated a strong preference for interactive learning methods such as discussions with colleagues, mentoring, and individually focused programming (Roberts, 2007). Though synergistic supervision was not specifically identified in this study, synergistic supervision meets each of the aforementioned criteria.

High quality supervision. The supervisory relationship was identified as a critical component of professional development because supervisors can provide ongoing challenge, assessment, and support (Armenio & Winston, 2001; Winston & Creamer, 1997; 1998). Arminio and Winston (2001) studied the behaviors, values, and attitudes that characterized a high quality supervisor. Quality supervisors aligned the staff through development and teamwork, reinforcing this process through high expectations and modeling the behavior they wanted to see from others (Arminio & Winston, 2001). Supervisors were clear about values, ethics, and principles of fairness; they interpreted and built upon the culture of the institution and had a vision of where the institution was going (Arminio & Winston, 2001). These expectations were conveyed through regular supervision meetings in both a group context (the entire staff) and individually. Staff

members were integrated into meeting planning processes; supervisors and supervisees engaged in much face-to-face interaction and this interaction was characterized by good communication skills. Furthermore, supervisors issued developmental challenges in manageable portions (Arminio & Winston, 2001). Arminio & Winston (2001) commented on the importance of soliciting information from multiple perspectives; the supervisor-supervisee dyad did not by itself constitute a sufficient feedback loop. They specifically noted the value of multi-rater instrumentation because this method can anonymously incorporate peer, subordinate, and superior perspectives into the supervisory process to provide a deeper context for developmental discussions (Arminio & Winston, 2001).

High quality supervisors were willing to confront subordinates when necessary, but these confrontations were couched within a systematic process of meetings (whether individual or group). Confrontation from high quality supervisors was not perceived to be micro-managing or characterized by overly intrusive methods. High quality supervisors struck a balance between providing structure and direction, while remaining open to subordinate interests. They made their points clearly and unobtrusively, but also listened to supervisees' perspectives. Supervisors communicated high levels of expectation, but also conveyed a sense of encouragement and support to supervisees. Arminio and Winston (2001) articulated this dynamic balance, "[Q]uality supervision is not controlling staff members, but rather setting the context, motivating, teaching, listening, observing, giving direction, and caring. The relationship is synergistic" (p. 41).

When supervisors were asked to describe how they managed to strike this synergistic balance, however, many confessed that they did not know or suggested they learned through trial and error. Some pointedly stated that they did not learn this skill in graduate school. Many said they learned by not doing the things that used to frustrate them as a supervisee. Given the importance of supervision, Armino and Winston (2001) claimed that more and better preparation was needed for supervisors to learn how to develop synergistic relationships. Toward this aim, they recommended better in-service education and training (Armino & Winston, 2001). Such training could be based around the framework of Winston and Creamer's (1997; 1998) synergistic supervision model for professionals. This model is an evolution of the one Winston, Ullom, & Werring (1984) proposed for RA supervision. Due to the greater responsibilities of professionals, this version offers a broader, more sophisticated process but the underlying principles are the same. Supervision is a synergistic balance between individual needs and desires and institutional objectives. The effective supervisor negotiates this balance through versatile interpersonal leadership skills and carefully developing and maintaining a high quality relationship with supervisees. Although Winston and Creamer's (1997; 1998) synergistic supervision model articulated the process for developing synergistic supervisors, Arminio & Winston (2001) claimed additional assessment and research were needed to advance the practice of synergistic supervision. The effective supervisors described by Arminio and Winston (2001) exhibited the characteristics of a versatile leader (Kaplan & Kaiser, 2006). An instrument such as the LVI-S can be a valuable tool to advance research on synergistic supervision and the training and development of synergistic supervisors.

Measuring the Development and Supervision of Student Affairs Professionals

Creamer and Janosik (2003) recommended a range of performance appraisal tools and processes used to measure performance in student affairs. They referenced Brown's (1988) model of quantitative performance appraisal—performance appraisal should be deliberate and collaboratively constructed. Assessment should focus on behaviors (Brown, 1988; Grote, 1996) and a number of different instrument designs were recommended. Brown (1988) strongly advocated for the behaviorally anchored scale. Such scales focus on a specific task or competency, e.g., “*Listens* effectively to others.” Then five to seven descriptions of listening behaviors are provided, ranging from *highly effective* to *poor*, and the rater selects which one is most characteristic of the ratee. An example of an effective listening behavior might state, “*Listens* intently, remaining focused on the communicator.” An example of poor listening behavior might state, “Frequently interrupts others when they are talking.” Better performing behaviors are ranked with higher values and lower performing behaviors are ranked with lower values (Creamer & Janosik, 2003). The limitation of this approach is the rater is unable to provide any indication of frequency. What if the supervisor was a good listener but was consistently unavailable? Grote (1996) recommended the use of frequency scales—an evolution of the behaviorally anchored scale. Rather than judging between good or bad behavior, this scale presents an idealized behavior and then asks how frequently it occurs (Creamer & Janosik, 2003). The authors intimated that behaviorally anchored scales and behavioral frequency scales were subject to limitations (Creamer & Janosik, 2003). Ratings could fall prey to bias, halo effect, and central tendency. Furthermore, poorly

constructed scales might measure personality characteristics in lieu of behavior. These scales are time-consuming and difficult to develop, but if time and effort is invested wisely, the feedback generated by such measures can be invaluable because they provide specific, behavior-based feedback that inform the developmental process (Creamer & Janosik, 2003).

One limitation not addressed by Creamer and Janosik (2003) was behaviorally anchored and behavioral frequency scales do not measure overuse of a skill or ability. Winston and Creamer (1997) alluded to overuse tendencies in their discussion of power, but the recommended measurement methods do not specifically account for overuse of power or overuse of other leadership characteristics. Synergistic supervision places great emphasis on the dynamic balance of behaviors. Kaplan and Kaiser's (2003a; 2006) research indicated that overuse of one type of skill will prompt under-use of that skill's complement. It seems natural, then, the measurement of synergistic supervision should incorporate a methodology that accounts for behavioral balance. The versatile leader model provides a behavioral assessment component that can complement and enhance the synergistic supervision model.

Summary of Paraprofessional and Professional Development, Supervision, and Leadership

Many of the interpersonal leadership themes and supervision models for RAs and professional staff are similar. Winston, Ullum, and Werring (1984) described a synergistic supervision model for RAs that emphasized the importance of attending to RAs interpersonal needs as well as to the goals and directives of the institution. Literature

on supervising and developing student affairs professionals has also been based on this synergistic supervision model—although suitably modified to apply to new professionals (Janosek, et al., 2003) and to experienced professionals (Winston & Creamer, 1997; 1998). Professional-level supervision and development is more comprehensive, but the fundamental interpersonal competencies of each are largely the same. Lombardo and Eichenger (2000) analyzed interpersonal leadership competency models from a broad selection of industries and reported 85% of interpersonal leadership competencies were the same regardless of profession. Paraprofessional roles in student affairs were not as broad or complex as professional roles, but the fundamental skills and challenges of both were highly similar. Eichinger and Lombardo (2000) reported that leadership competency models from diverse industries exhibited 85% overlap. It stands to reason paraprofessional and professional skills in student affairs would prove highly similar. Therefore it appears that much of the interpersonal leadership research and guidelines that have been proposed for paraprofessionals are equally valid for professionals, and vice-versa. Furthermore, since many Master's-level HDs are studying to become professionals in student affairs, adopting measurement processes and personal development plans that are similar to professional-level programs can enhance the professional development value of the HD role. This value is further enhanced through the self-awareness and self-developmetn that can result from this practice.

Contemporary Leadership Models in Residence Hall Leadership: Theory and Measurement

Residence hall leadership is situational in nature; a HD or RA might excel in one situation but struggle in another (Upcraft & Pilato, 1982; Winston & Fitch, 1993). There is no predictability or consistency for the types of situations that occur in a residence hall environment, therefore RAs and HDs must be ready to deal with situations that range from the frivolous to the life threatening (Upcraft & Pilato). Because the type of leadership style required depends on the situation, Upcraft and Pilato offered six key questions to ensure leaders have thoroughly processed the situation before they attempt to take action: 1) Do I have all the facts I need? 2) Do I have conceptual understanding of the situation? 3) Do others confirm my perception of the situation? 4) Have I faced similar situations in the past, and if so, what did I learn? 5) Who are the people involved in the situation? 6) What are my predictions as to the situation's outcome? These questions help the leader tactically assess the situation, but these questions do not help an RA decide how to take action. Though not explicitly stated, these suggestions are reminiscent of the strategic intent of Vroom and Yetton's (1973) decision-making model—help leaders broaden the criteria they use when making decisions.

Situational Leadership

To train residence hall leaders on how to match behavioral interventions with the needs of a situation, Upcraft and Pilato (1982) recommended using Hersey and Blanchard's (1977) model of situational leadership. Using the situational model, in theory, permits a leader to match the appropriate leadership style with the needs of a

situation. Because of its prescriptive nature and because it makes intuitive sense, the situational model is popular in training and development programs (Northouse, 2004).

The critical literature on situational leadership, however, reveals significant problems with the theory. Very few published studies validate its theoretical foundation (Northouse, 2004). Situational leadership has vaguely defined constructs (Graeff, 1997; Vecchio, 2007; Yukl, 1989) and studies by Vecchio (1987), Fernandez and Vecchio (1997), Norris and Vecchio (1992), and Vecchio and Boatwright (2002) produced results that could not be explained by the model (Northouse, 2004; Vecchio, 2007). Northouse (2004) reported Graeff (1983) and Yukl (1989) criticized the questionnaires used in situational leadership because response options were limited to leadership styles described within the situational leadership framework. Other leadership behaviors are not acknowledged nor accounted for in the instrumentation. A reliable and valid assessment instrument has not yet been provided for measuring situational leadership (Vecchio, 2007).

In spite of measurement limitations, Hersey and Blanchard's (1977) situational leadership model provided a valuable contribution to leader training and development because it emphasized the importance of situational assessment and behavioral flexibility (Bass, 1990). The popularity of Hersey and Blanchard's model underscored the demand for intuitive, behavior-focused leadership assessment and training tools.

Transformational Leadership

The transformational leadership model has been highly regarded in student development circles because the values and tenets of transformational leadership align

closely with those of higher education institutions (Astin & Astin, 2000; Dugan, 2006; Komives, Lucas, & McMahon, 2007; Komives, 1991a; 1991b; 1991c). Transformational leaders create a transcendent spirit among their following and challenge them to expand their conceptualization of problems and challenges (Avolio, Bass, & Jung, 1999). Followers feel connected to the mission of the organization, operate from a sense of intrinsic rather than extrinsic gratification (Kouzes & Posner, 2004), and believe that their accomplishments benefit the greater good of society. Burns (1978) stated that transforming leaders engage constituents in an interdependent relationship based on shared motives, values, and goals; effective leaders create meaningful social change that satisfies the needs and expectations of stakeholders. This egalitarian approach toward leadership suits an increasingly complex society where leaders must learn to negotiate according to shared needs rather than from positions of force or power (Burns, 1978; Dreikurs, 1971; Friedman, 2007).

Transformational leadership according to Bass and Avolio. Transformational leadership differs from behavioral and situational models because it addresses concepts such as charisma, inspiration, and intellectual stimulation. According to Bass and Avolio (1994), leaders engage in transformational leadership, transactional leadership, and laissez-faire leadership (non-leadership). Transformational factors include idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration (Bass & Avolio, 1994). Idealized influence is the degree to which leaders model those behaviors followers would emulate. Inspirational motivation addresses the leader's communication of a vision and the articulation of high aspirations and

expectations. Intellectual stimulation refers to the leader's penchant for challenging followers to think critically and differently about their beliefs. The final factor, individualized consideration, is the degree to which a leader makes an effort to listen to the individual needs of followers (Bass & Avolio, 1994).

Contingent reward and management-by-exception are transactional factors within the model. Contingent reward behaviors establish transaction-based agreements through outlining task-related performance expectations for followers; then when followers meet expectations, they are rewarded for their effort (Bass & Avolio, 1994). Management-by-exception is composed of punitive behaviors. If expectations are not met, then negative reinforcement, constructive criticism, or negative feedback loops ensue (Bass & Avolio, 1994). Laissez-faire leadership is a non-leadership factor. The laissez-faire leader may not be present (e.g., the absentee landlord) or is disengaged to the point where their presence is not felt or recognized (Bass & Avolio, 1994).

The transformational leadership model provided a different and expanded way of conceptualizing leadership behaviors. Transformational leadership outlines connections between a leader's aspirations and values and those of followers. Transformational leadership attracted the attention of colleges and universities, becoming one of the first leadership theories empirically tested in a residence hall environment. Data gathered from this assessment process was reported to inform RA and HD leadership development programs (Komives, 1991b; 1991c; Posner & Brodsky, 1993).

Assessing leadership using Bass and Avolio's transformational leadership model. Transformational leadership is most frequently measured through the Multifactor

Leadership Questionnaire (MLQ) (Lowe, Kroeck, & Sivasubramaniam, 1996). The MLQ is a multi-rater instrument developed for professional organizational applications such as the military, religious and educational institutions, and public and private businesses (Lowe, Kroeck, & Sivasabramaminiam, 1996). The instrument was originally conceived by Bass (1985), but has been updated on multiple occasions (Bass & Avolio, 1995/2000). Although the MLQ has been used with traditional-age undergraduates in select cases (Komives, 1991a; 1991c), more often than not, “students” referenced in MLQ research are actually working students with a mean age greater than traditional undergraduates (e.g., Tepper & Percy, 1994) or students who are MBA graduate students.

A number of studies have investigated the validity of the MLQ. Several researchers have challenged the proposed factor structure of the model (Bycio, Hackett, & Allen, 1985; Tejada, Scandura & Pillai, 2001; Tepper & Percy, 1994). These studies questioned high levels of intercorrelation between transformational factors and suggested that Bass and Avolio might revise the model into a simpler form. Other researchers challenged those findings, indicating that although intercorrelation was present, the effects were negligible and were characteristic of latent, interdependent factors (Avolio, Bass & Jung, 1999; Antonaikis, Avolio, & Sivasubramaniaim, 2003). Several meta-analyses of transformational leadership have supported its validity (Lowe, Kroeck, & Sivasubraminaiam, 1996; Bono & Judge, 2004) and transformational-based leadership programs have been found to yield significant results that support maintaining the more complex structural model (Parry, 2005).

The MLQ is significant with regard to residence life leadership development because Komives (1991a; 1991b; 1991c) was the first to publish a study utilizing an empirically sound, multi-rater instrument to assess HD leadership styles and effectiveness. Komives (1991b) noted that a limitation of the MLQ is that it was designed for professional managers operating in work contexts quite different from that of a residence-life organization (Komives, 1991b). In the intervening years, there does not appear to have been a comparable multi-rater study published that focused on the HD/RA supervisory relationship as a leadership research context.

Housing and residence life studies using the Multifactor Leadership

Questionnaire.

Transformational factors, RA satisfaction, motivation, and HD effectiveness.

Komives (1991b) examined the relationship between HD self-assessed interpersonal leadership styles (transformational versus transactional), RA levels of satisfaction and motivation, and RA perceptions of HD leadership style and their effectiveness as a leader. The sample included 84 HDs and 806 RAs from seven public university residence life programs.

The most effective HDs emphasized an engaged, transformational, leadership style characterized by high consideration for others, charisma, and intellectual stimulation (Komives, 1991b). Hall directors also engaged in transactional leadership (Komives, 1991b). This style is characterized by less interpersonal engagement, constructing transactional agreements contingent on an RA's achievement of performance goals or objectives, and leaders who typically do not intervene unless there is a problem requiring

attention (Bass, 1990). Given the stressful nature of resident advising and the priority that RAs place on supportive relationships, the preference for more engaged, transformational leadership styles was not surprising (Komives, 1991b).

Self versus observer ratings comparisons indicated potential blind spots for HDs. In comparison to RA perceptions, HDs overrated themselves on the amount of positive, highly engaged leadership behaviors they provided. HDs underrated themselves on less engaged leadership characteristics (e.g., remaining distant unless a problem occurred) or non-leadership behaviors (Komives, 1991b). In short, HDs exhibited the similar perceptual bias characteristics as leaders in other organizational settings (Harris & Schaubroeck, 1988; Conway & Huffcut, 1997).

HDs also overestimated RA satisfaction and motivation levels in comparison to levels reported by RAs (Komives, 1991b). This disparity between self-report and other-report has been referred to as the false consensus bias (Krueger & Clement, 1994), and studies within business and industry have demonstrated similar patterns (Conway & Huffcut, 1997; Lombardo & Eichinger, 2000). It seems plausible that the method through which leadership data is collected (i.e., self-report versus other-report) is critical. Komives' (1991b) findings with HDs supported previous recommendations related to residence hall leader performance assessment: to obtain a fair and balanced evaluation, attempts should be made to access information from multiple sources to gain a more comprehensive perspective of the leader's performance (Upcraft & Pilato, 1982; Winston & Fitch, 1993).

Transformational factors and achievement styles. Komives (1991a) studied HD achieving styles and how they related to transformational and transactional leadership characteristics. Achieving styles were measured through the L-BLA Achieving Styles Inventory (Lipman-Blumen & Leavitt, 1979). Lipman-Blumen and Leavitt's (1979) achieving style theory posited that individuals have preferences for how they choose to complete tasks. They divided these styles into three broad categories: direct, instrumental, and relational. A direct achieving style is individualistic. The direct individual will identify personal goals and individually achieve those goals. The instrumental achiever sets personal goals, but attains these goals through positional influence (e.g., title or status), or through the actions of others. Relational achievers are other-focused; goals are set collaboratively and are achieved through collective effort.

Leadership constructs were measured through the Multifactor Leadership Questionnaire Form 5 Revised (Bass, 1985). A multi-rater methodology was used; HDs rated themselves on achieving styles and transformational and transactional leadership qualities. In addition, RAs provided their perceptions of HDs on these same constructs. Analysis was conducted on three levels: individual level (self-reported HD ratings), other level (RA ratings of HD), and work group level (responses from the HD and at least 50% of their RAs). Participants included 74 HDs, 602 RAs, and responses were distributed such that 64 work groups could be analyzed.

In the self-reported leadership responses, men and women reported similar perspectives on their leadership. Both ranked themselves highest on individualized consideration, then intellectual stimulation, charisma (this variable was renamed

idealized influence in later versions of the MLQ), contingent reward, and finally inspirational. The only significant difference between male and female self-reported scores was that males reported higher scores for intellectual stimulation.

In the self-reported achieving style data, HDs reported the relational achieving style as their most preferred; there were no significant differences between male and female mean scores on the relational scale. The direct achieving style was the second most preferred approach; but in contrast to the relational style, there was a significant gender difference: males scored significantly higher. Both men and women ranked the instrumental style as their least preferred, but like the direct style, men scored significantly higher on the overall scale scores for instrumental achieving.

Gender differences in achieving style and perceptions of transformational leadership became more pronounced when Komives correlated achieving style scales with transformational leadership scales. Male HDs who rated themselves high on transformational characteristics attributed achievement to their ability to take charge and to direct or control others. Relational achievement did not factor into their personal views of transformational leadership. In fact, the only significant correlation on relational achieving styles was inverse—higher scores on intellectual stimulation indicated lower reported preference for the relational achievement style. Female HDs held a dramatically different perspective. It appeared that female HDs believed their relational achievement style contributed to transformational leadership. All female relational achievement subscales were significantly and positively correlated with transformational scale scores.

When HDs' leadership and achievement styles were rated by RA observers, a common pattern emerged for both males and females: all HDs who were perceived to be transformational used collaborative, relational leadership styles. This perception affirmed the female perspective on leadership styles but contrasted sharply with Male HD self-reported achievement and leadership styles (Komives, 1991a). Komives further indicated that male supervisors should reconsider the belief that power-oriented or direct achieving styles lead to transformational leadership. Males should recognize that subordinates perceived their relational behaviors as having more transformational influence than other achieving styles.

Transformational factors, gender differences and RA satisfaction. Researchers have considered how factors such as gender might influence the leadership process between HDs and RAs. For example, Komives (1991c) explored the degree to which gender pairings of RAs and HDs affected RAs level of satisfaction with HD leadership, overall job satisfaction, and motivation. Gender pairings were not found to significantly impact leadership outcomes in any of the four combinations studied (Komives, 1991c). Komives (1991c) observed that these findings supported previous studies by Bartol and Wortman (1976) and Stitt, Schmidt, Price, and Kipness (1983). Since gender does not appear to have a significant impact on RA residence hall leadership, gender differences will not be assessed in this study.

The Five Practices Model

The five practices of exemplary leadership is another model of transformational leadership (Sashkin, 2004). The authors collected case studies from over 1200 professional managers, asking them to report their “personal best” leadership experiences. These accounts were content-analyzed to discern specific leadership characteristics from each case. Using these characteristics, Kouzes and Posner (1987) developed an extensive list of questions about leadership behavior. Hundreds of managers were asked to answer these questions by describing exceptional managers they knew or with whom they had worked previously. After analyzing the results, five clearly defined factors emerged. Each factor could be described in specific, behavioral terms. Using behavioral descriptors based on these factors, Kouzes and Posner created a multi-rater instrument named the Leadership Practices Inventory (LPI). The five variables measured by the LPI are described as follows:

1. “Modeling the way”—leaders set an example for others to follow. Leaders also provide guidance on how to accomplish goals through step-wise processes, breaking down large goals into more manageable and attainable processes.
2. “Inspiring a shared vision”—the actions a leader takes to build and articulate a vision of the future and engender support for this vision.
3. “Enabling others to act”—actions leaders take to enable others to manage tasks; for example, leaders build collaborative rather than competitive

networks; leaders support followers as they engage challenges and opportunities for personal growth.

4. “Challenging the process”— behaviors a leader uses to encourage others to engage in calculated risk-taking, innovate new solutions, and otherwise improve the organization.
5. “Encouraging the heart”—the encouraging leader celebrates group accomplishments and individually recognizes followers’ contributions contributing to group accomplishments.

Kouzes and Posner’s (1987) model was empirically derived, and this concrete evidentiary base lends credibility to the model (Sashkin, 2004). The items also are more specific and behaviorally focused than those of the MLQ (Sashkin, 2004), so feedback can target specific behaviors that contribute to effective leadership. Using the same instrument development methods as Kouzes and Posner (1987), Brodsky (1988) designed a student-focused version of the LPI. Results from both the student study and the professional study pointed toward the same five factors. These five factors were crafted into a leadership model called the five practices of exemplary leadership. *The Leadership Challenge* text, training curriculum, and the Leadership Practices Inventory (SLPI) were founded on this competency model.

Measurement of Kouzes and Posner’s model. The SLPI is one of very few leadership instruments designed for college students that has been validated within a college context (Posner, 2004; Schwartz & Gimbel, 2000). Also, the SLPI is the only multi-rater leadership assessment instrument for students (Leadership Challenge, 2008).

The SLPI is a brief, 30-item instrument that has a “self” form for the focal leader to complete and an “observer” form for raters to complete. Participants score their perceptions of a leader’s behavior on five distinct “practices” of leadership: Modeling the Way, Inspiring a Shared Vision, Enabling Others to Act, Challenging the Process, and Encouraging the Heart. Each practice variable is measured with 6 items, and each item is measured with a response scale that ranges from 1 (*rarely observed*) to 5 (*frequently observed*). Numerous studies have demonstrated the usefulness of such an instrument in a variety of college and university contexts. The SLPI has been used to study leadership in fraternities and sororities (Posner & Brodsky, 1992; Posner & Brodsky, 1994; Posner, 2004), among RAs (Posner & Brodsky, 1993), orientation advisors (Posner & Rosenberger, 1997), and ROTC students (Baxter, 2001).

The SLPI also has weaknesses. The response format permits raters to indicate frequency of behaviors observed, but raters are unable to indicate their perceptions of the leader’s effectiveness. Although frequency scales are the most popular type of scale used in multi-rater assessment (Leslie & Fleenor, 1998), Shipper (1991) reported that increased behavioral frequency might exist independent of behavioral mastery. Yukl (1989) recommended further study examining the quality of leadership behaviors, not merely the quantity of behaviors. The SLPI has limited value for analyzing quality of behaviors because it does not incorporate an evaluation scale or open-ended questions.

Five Practices and RA leadership and effectiveness. Posner and Brodsky (1993) studied RA self-perceptions of leadership, resident perceptions of RA leadership, and leadership effectiveness as evaluated from multiple perspectives. These perspectives

included the RA's self-rating, observation from residents, and from RA supervisors. Posner and Brodsky sampled 333 RAs and each RA was rated by an average of 3.9 residents for a total of 1304 observers. RAs and resident observers completed the SLPI to measure leadership characteristics, a 9-item effectiveness questionnaire designed by the authors, and a demographic questionnaire to collect about participant profiles. Supervisors were asked to provide a separate and independent global effectiveness rating of RA performance; they were to consider all of their RAs as a group, then assign a (1) to RAs who were among the least effective performers and assign a (3) to RAs who were among the most effective performers. All other RAs were automatically given a rating of (2) to represent average performance. This procedure divided RAs into high, average, and low performance quartiles.

Posner and Brodsky reported RA perceptions of leadership and leadership effectiveness paralleled those of their residents. RAs who stated they engaged in the five practices more frequently also viewed themselves as more effective leaders. Correspondingly, residents who rated RAs with higher scale scores also perceived these RAs as being more effective than those with lower scores. The relationship between effectiveness and leadership scores remained consistent across all RA and constituent-rated performance categories. High performers scored highest on the practices; average performers scored higher than the low performers, but not as high as the high performers; and low performers rated the lowest scores. According to Posner and Brodsky (1993), "Indeed, what may set apart the most effective RAs from their counterparts is their above-average use of these leadership practices" (p. 303).

Supervisor ratings of effectiveness revealed a different relationship between effectiveness and leadership. Supervisor data showed that RAs scoring high on encouraging, modeling, and enabling were rated as more effective than RAs who reported engaging in these behaviors to a lesser degree. But unlike resident and RA effectiveness ratings, challenging the process and inspiring a shared vision behaviors did not contribute significantly to supervisor effectiveness ratings.

Two possible explanations were provided for this finding. First, challenging and inspiring behaviors may be more difficult to observe because (a) they may not be visible or tangible, especially given the lower frequency of interactions between RAs and supervisors in comparison to RAs and residents; (b) RAs might have self-monitored these behaviors in front of supervisors because it was not politically savvy to let them know the frequency to which they experimented with new ways of doing things. Another possible explanation, at least with regard to challenging the process, was this behavior might not be relevant to the job performance and success of RAs. Supervisors might view challenging behaviors as contraindicative of good leadership in residence hall settings. Posner and Brodsky stated, “those RAs who challenge the process may be viewed unfavorably by their supervisors when it comes to enforcing and acting to university rules and regulations” (p. 304).

With respect to challenging the process skills, were effectiveness ratings inconclusive because of inappropriate use of the skill or was it because of a perceived absence of the skill? The research design could not provide a clear indication, and thus limits the effectiveness of feedback because of increased attributional ambiguity (Kruger

& Dunning, 1999). Posner and Brodsky (1993) alluded to this limitation when they stated that future research might, “secure assessments from the Resident Directors of the extent to which they perceive the RAs engaging in the various leadership practices, and not just the extent to which they were performing well in their positions” (p. 304). The authors were calling for researchers to incorporate both a measure of frequency and a measure of effectiveness to gain better perspective of the relationship between leadership behaviors and effective outcomes. This suggestion was consistent with Yukl’s (1989) recommendation that leadership research should study both frequency of and quality of leadership behaviors.

Effectiveness as a Criterion Variable in Residence Hall Multi-rater Studies

Reviewing the greater body of leadership literature, effectiveness has been evaluated through multiple methods. Such methods include effectiveness questionnaires (Brodsky & Posner, 1993; Tsui, 1984; 1986) through organizational performance outcome data such as a manager’s percent-to-goal or supervisor-assessed performance reviews (Lowe, Kroeck, & Sivasubramaniam, 1996), and through subjective assessments provided by a panel of subject matter experts (Jager & Ciason, 2005). In the residence life leadership literature, where studies encompass multiple campuses, the questionnaire-based method has been the assessment method of choice (Komives, 1991b; Posner & Brodsky, 1993).

Questionnaire-based effectiveness measures and residence halls. Effectiveness questionnaires are usually rated with a Likert-type response scale designed to gauge the level of performance of a leader. Podsakoff and Organ (1986) noted that caution should

be used when measuring leadership skills and effectiveness through common methods, such as using a leadership questionnaire and an effectiveness questionnaire. They provided recommendations to help control for the limitations of common method variance (Podsakoff & Organ, 1986). Effectiveness measures and leadership measures can be integrated in a variety of ways. Some leadership measures, such as the MLQ, incorporate an effectiveness scale into the instrument (Bass & Avolio, 1985; 1998; Lowe, Kroeck, Sivasubramaniam, 1996). Komives (1991b) utilized the effectiveness scale embedded within the MLQ. Some effectiveness scales may be used independent of other instrumentation. Tsui (1984) created an effectiveness scale based upon reputational effectiveness. Tsui's three-item effectiveness scale measures how well a leader has met the observer's performance expectations. Sometimes researchers may also design context-specific measures to complement the leadership instruments used in their study. Brodsky and Posner (1993) adopted this approach in their study of RA leadership practices.

Meta-analysis of Multifactor Leadership Questionnaire research. In a meta-analytic study of the MLQ and leadership effectiveness, Lowe, Kroeck, and Sivasubramaniam, (1996) hypothesized that the type of criterion variable would moderate the relationship between type of leadership style and effectiveness. They distinguished between two types of criterion variables for effectiveness: effectiveness as rated by subordinates (the embedded MLQ effectiveness scale) and effectiveness as rated by organizational measures (e.g., profit, percent of goals met, or records such as supervisory performance appraisals).

The hypothesis was fully supported. Subordinate perceptions of effectiveness yielded significantly higher positive relationships in comparison to organizational effectiveness measures. Lowe, Kroeck, et al., (1996) observed that common source variance contributed to inflated correlations because raters strived to obtain consistency in their responses to the dependent and independent variables—a common sub-category of common method variance as reported by Podsakoff and Organ (1986). Lowe, Kroeck, et al., suggested that organizational criteria probably attenuated the relationships between leader behavior and effectiveness because such measures focus the dependent variable on a narrow perspective of performance that may not account for the full spectrum of perceptions that might be relevant to assessing leadership effectiveness. Therefore the most reasonable estimate of effectiveness was probably somewhere in between the inflated results from the embedded scale and the under-represented results from the organizational measures.

In the methods section of her study on hall directors and selected resident advisor outcomes, Komives utilized the effectiveness scale incorporated within the MLQ (Komives, 1991b). Use of this embedded scale exposed the study to common method variance. Komives did not identify common method variance in the limitations for her study, however. Results demonstrated significant relationships between effectiveness and transformational behaviors, but these results should be interpreted with this limitation in mind.

Distinct rater groups in HRL research. The research design proffered by Posner and Brodsky (1993) was notable for several reasons. Tsui (1984) and Conway and

Huffcut (1997) noted that distinct rating groups contribute different perspectives on leader performance. Distinct rater groups have unique opportunities for observation. A resident might observe an RA more frequently than a supervisor and these observations take place under different circumstances. Therefore, residents observe a different aspect of the RA's leadership than does a supervisor (Posner & Brodsky, 1993). Furthermore, the behaviors a resident might associate with effectiveness may or may not be the same as a supervisor (Conway & Huffcut, 1997). For example, a supervisor might consider the leader's ability to maintain accurate records just as important as the RA's interpersonal skills. If the RA gets along well with constituents but does not manage organizational responsibilities, then the supervisor's effectiveness ratings might indicate this discrepancy whereas a resident's ratings probably would not.

In addition to providing a broader observational spectrum, utilizing multiple sources helps to control for common method variance (Podsakoff & Organ, 1986). Posner and Brodsky requested that supervisors complete a global effectiveness rating rather than the same effectiveness scale used by RAs and their residents. This approach altered both the type of data collected (global effectiveness versus ratings of specific behavioral effectiveness) and provided an alternate rating source (supervisor effectiveness ratings versus resident effectiveness ratings).

Summary

Good leadership benefits countries, communities, and organizations; bad leadership harms them. Astin and Astin (2000) argued the quality of leadership in the United States is in decline. Given the state of the national economy, tepid social capital

levels, and serious demographic challenges, this point is tough to argue. The challenges of leadership are growing more complex but the number of individuals capable of facing these challenges is shrinking. Leadership development, particularly for younger leaders who will take over as Baby Boomers retire, is a critical need. These individuals will face more complex challenges than their predecessors, and will do so with less experience.

Leadership development technology has improved greatly over the last three decades. Research methods and analytical processes have advanced. Improved instrumentation contributed to these gains as well. The science of multi-rater measurement has helped researchers model leadership phenomena with greater precision. Kaplan and Kaiser's (2006) versatile leader model and the Leadership Versatility Index are examples of these technological improvements.

The Leadership Versatility Index helps to provide bridge between the complex nature of leadership and the use of competencies as a development strategy. Prior to the LVI, instruments were unable to measure the dynamic balance of behaviors because they could not measure excessive use of strengths. The instrument's innovative rating scale and the duality-based design enable it to measure balance reliably and with validity. Leadership versatility, as measured by the LVI, explained half the variance of executive effectiveness.

The LVI was designed for an executive-level population. Because development of younger, less experienced leaders is imperative, it makes sense to modify the LVI for use with this younger population. The entire instrument, however, does not need to be modified. The forceful/enabling duality focuses on the interpersonal processes leaders

use to influence others—it measures how people lead. These interpersonal processes are as relevant to college student development as they are to executive development. The forceful/enabling duality will be modified and the more executive-appropriate, strategic/operational duality will be discarded from the modified version of this instrument.

Student leadership research has not kept pace with professional research. Research methods are less sophisticated and analytical methods have lagged behind those of professional leadership research. The overdependence on self-report data collection and failure to control for common method bias are two examples of these shortcomings. Applying multi-rater research methods and analyzing results at different levels could reveal context-specific results about where and how students develop leadership skills—a question that has yet to be answered.

The context for student leadership research is important. The housing and residence life context is well suited for leadership research. The formal leader-member relationship between hall directors and their resident advisors is also advantageous for a multi-rater validation study. Supervision styles have received some attention in housing and residence life literature. The synergistic supervision model recommends that supervisors maintain a relationship that balances between being open and focused on the supervisee, but also keeps the institution's interests in focus as well. The Leadership Versatility Index is an excellent fit with this model. Both emphasize the importance of balance and of the paradoxical nature of leadership.

The last published multi-rater study incorporating hall directors and resident advisors was conducted in 1991. It makes sense to revisit this research context since much has been learned about residence hall leadership, about leadership assessment and measurement since the previous study. The Leadership Practices Inventory—Student has been used in the residence hall context with some success before. Since it is the only student focused, multi-rater instrument that has received psychometric scrutiny, the Leadership Practices Inventory-Student is the preferred instrument to use for content validity purposes.

CHAPTER III

METHODOLOGY

Chapters I and II presented the rationale and literary support for the development of a campus focused, multi-rater assessment instrument named the Leadership Versatility Index—Student (LVI-S). A review of student leadership literature demonstrated the need for multi-rater instruments that are capable of assessing Hall Director (HD) and Resident Advisor/Assistant (RA) relationships. This review also demonstrated the need for an instrument that is able to measure leadership strengths, shortcomings, and strengths overused. In this chapter, the methodology for developing and validating a Leadership Versatility Index for Students is explained. Research questions and hypotheses are presented, participants and instrumentation are described, and an overview of the procedures for data collection and statistical analyses is provided.

Research Questions and Hypotheses

The present study will explore the relationship between the dimensions of versatile leadership, HD performance, and RA ratings leadership effectiveness. A series of reliability and validity analyses will be conducted to determine the psychometric properties of the LVI-S.

Research Question 1: Does the LVI-S verify that there are two leadership factors: forceful leadership and enabling leadership?

Hypothesis 1: The LVI-S will present a clearly identifiable two-factor structure.

These two factors will have an inverse relationship to one another.

Research Question 2: What proportion of the HD reputational effectiveness can be explained by the six LVI-S sub-dimensions? Are any of the six sub-dimensions statistically significant predictors of HD reputational effectiveness as reported by RAs?

Hypotheses 2a: Scores associated with the items for each sub-dimension (i.e., HD sub-dimension scores, as reported by RAs), will be statistically significant predictors of leader effectiveness ratings.

Hypothesis 2b: HD self-reported sub-dimension scores will not be statistically significant predictors of leader effectiveness ratings.

Research Question 3: To what extent do scores on the LVI-S and the SLPI correlate, providing evidence of convergent construct validity of the LVI-S?

Hypotheses 3a: LVI-S sub-dimensions of Enabling leadership will significantly correlate with Enabling Others to Act, Inspiring a Shared Vision, and Encouraging the Heart variables on the SLPI.

Hypothesis 3b: LVI-S sub-dimensions of Forceful leadership will significantly correlate with the Challenging the Process variable on the SLPI.

Research Question 4: Do most HDs underestimate or overestimate their F/E dimension scores in comparison to RAs ratings of the HD's F/E dimension scores?

Hypothesis 4: Most HDs will overestimate their leadership scores on the six F/E sub-dimensions (rate themselves closer to the right amount) than the Resident Advisors' ratings of HD.

Research Question 5: Is there a statistically significant mean difference in effectiveness ratings between HDs who overestimate their F/E dimension scores versus HDs who underestimate their F/E dimension scores?

Hypothesis 5(a): There will be a statistically significant mean difference in effectiveness ratings between HDs who overestimate their F/E leadership sub-dimension scores versus those who underestimate their F/E leadership sub-dimension scores.

Hypothesis 5(b) HDs who overestimate their F/E leadership sub-dimensions scores will be rated less effective by RAs than HDs that underestimate their F/E leadership sub-dimension scores.

Research Question 6: Which regression model better predicts reputational effectiveness scores: Model 1, an additive model where interaction between complementary F/E sub-dimensions is not accounted for, or Model 2, a model that incorporates the focal leader's joint standing scores (versatility) on complementary F/E sub-dimensions?

Hypothesis 6: Model 2 will explain a greater percent of the variance in reputational effectiveness; model 2 will more effectively predict reputational effectiveness scores.

Participants

There are two primary categories of participants responding to this multi-rater study: Hall Directors ($n = 57$) and Resident Advisors ($n = 262$). Hall Directors were the focal leaders of the study; Resident Advisors (RAs) currently being supervised by the HD provided observer ratings. All participants were actively serving as a Hall Director (HD),

or as a Resident Advisor (RA) currently being supervised by the focal leader. Data was collected from seven colleges and universities in the Southeast United States. A summary of demographics and categorical characteristics has been provided in Chapter 4. The total number of participants in the sample met the calculated targets. Using a power calculation, a sample size of 52 leaders was need to obtain .90 power at alpha < .05 (Cohen, 1988). In addition, the proposed analyses required an average of approximately four raters per leader; the observed raters-to-leader ratio was 4.9 to 1.

Instrumentation

Participants completed two leadership instruments, an effectiveness scale, and a brief demographic questionnaire. The order of the instruments was as follows: Demographic Questionnaire, LVI-S, SLPI, and Tsui Reputational Effectiveness Scale. Below, the development and psychometric properties of each instrument are described. A copy of instrumentation is included in Appendices A-D.

Leadership Versatility Index-Student (LVI-S)

LVI-S feedback is captured through a computer-based form consisting of twenty-four items. Raters use a bidirectional scale to score behaviors on a continuum ranging from -4 (*much too little*) to 0 (*the right amount*) to +4 (*much too much*). These twenty-four items define the highest level of the LVI-S hierarchy, the Forceful / Enabling duality. This duality is divided into two symmetrical dimensions, the Forceful dimension and the Enabling dimension.

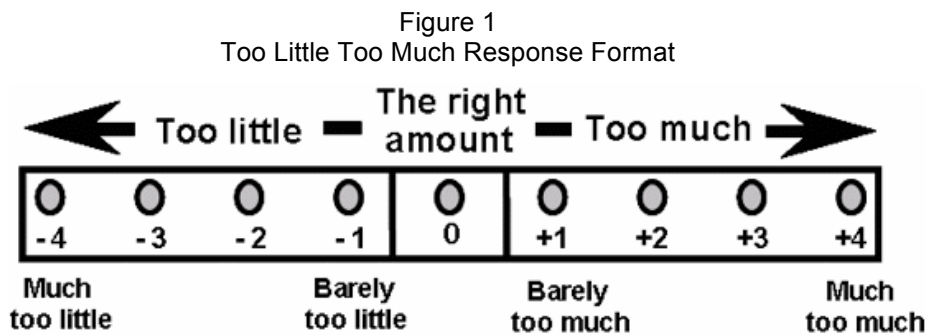
Each side of the Forceful / Enabling duality complements and opposes the other. Forceful and Enabling dimensions are divided into three sub-dimensions consisting of

four items each. These divisions are symmetrical, such that a 4-item sub-dimension on the Forceful side mirrors a 4-item sub-dimension on the Enabling side. This structure forms three complementary sub-dimension sets. Each set references a specific perspective of leadership. The first set, *Takes Charge / Empowers Others*, addresses locus of power. *Takes Charge* measures the actions leaders take when they assert authority; *Empowers Others* measures the behaviors used to imbue others with leadership responsibility. The second set, *Declares / Listens*, addresses decision-making. The *Declares* sub-dimension assesses leaders' use of autocratic decision-making strategies; *Listens* measures leaders' use of democratic decision-making strategies. The third set of sub-dimensions, *Pushes / Supports*, addresses leaders' motivational orientation. Behaviors within the *Pushes* sub-dimension focus on task-completion, performance and exhibit low concern for people. Conversely, behaviors within the *Supports* sub-dimension focus on understanding others' perspectives, expressing empathy, and encouragement.

Overall, the Forceful dimension consists of items 1-12 and the Enabling dimension consists of items 13-24. Opposite sides of the duality are joined through complementary, pair-wise relationships between items. The first item in the Forceful dimension, ([Ftc_1] Forceful, *Takes Charge*, item 1), is the complementary opposite of the first item in the enabling dimension ([Ee_13] Enabling, *Empowers Others*, item 13). Item Ftc_2 serves as the complementary opposite for item Ee_14, etc. The complementary opposite relationships established at the item level, may be aggregated throughout the hierarchical structure of the LVI-S, thereby defining paradoxical relationships between sub-dimension sets as well as between dimensions.

Leadership versatility is calculated by measuring the relationship between complementary opposites on either side of the Forceful / Enabling duality. A joint-standing score computes the leader’s relative standing on complementary items, sub-dimensions, and dimensions. This score provides a joint comparison between values on one side of the duality with their matching complementary values on the other side.

The most discrete level of measurement occurs between complementary item pairs. Responses to each item in a pair are recorded through the Too Little Too Much (TLTM) response format (Figure 1). Values on the TLTM format range from -4 to +4, and the ideal score is the central value, 0, designated as “*the right amount.*”



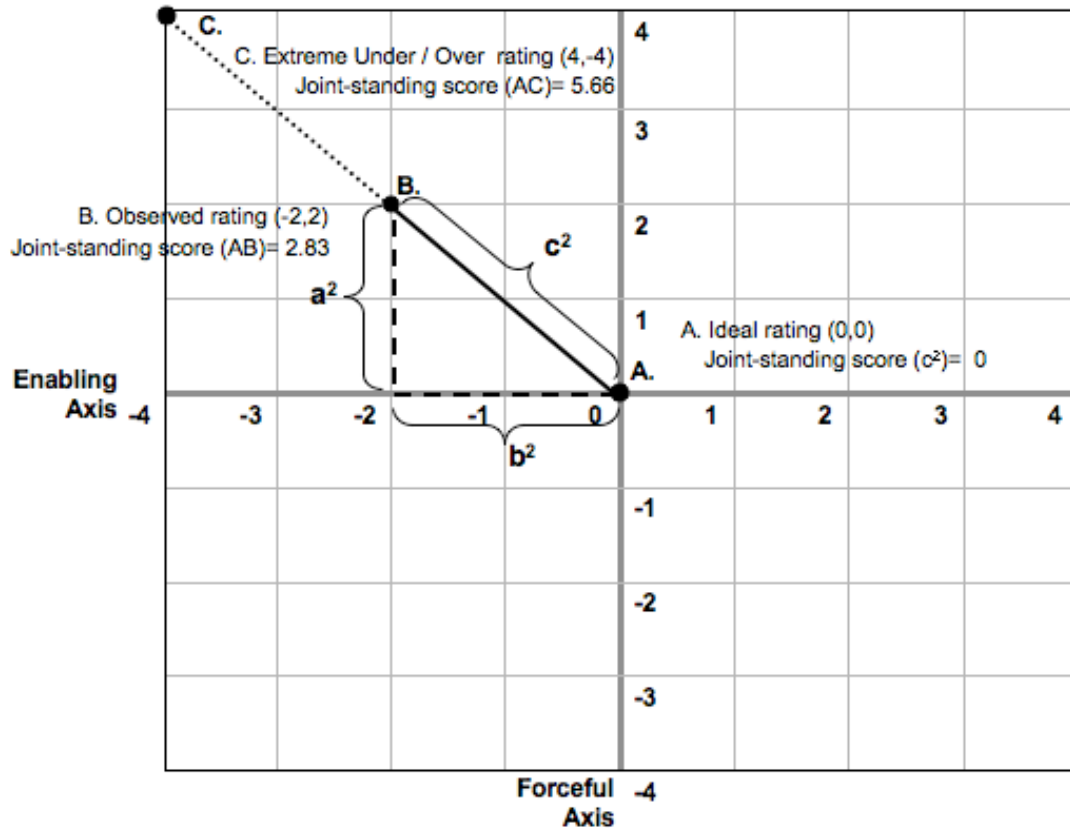
The joint relationship of item pairs may be graphically illustrated by turning response scales for complementary items 90 degrees to one another. This forms a graph where the X and Y-axes range from -4 to +4, and intersect at the origin (0, 0) (see Figure 2).

Observers’ responses may be plotted as a coordinate pair (an Enabling score plots on the X-axis and a Forceful score plots on the Y-axis).

The joint-standing score is calculated through the Pythagorean theorem ($a^2 + b^2 = c^2$). The ideal score on a matched pair of Forceful and Enabling items would be plotted at

(0, 0), indicating both items were scored at “*the right amount*” (see Figure 2, point A). Inserting these ratings into the Pythagorean formula produces a joint-standing score of zero, because $0^2 + 0^2 = 0^2$. This joint-standing score represents perfect versatility between the two items. Conversely, a score of (-4, 4), or any permutation of extreme positive or negative ratings, produces the least favorable joint-standing score (see Figure 1, point C). When inserted into the Pythagorean formula, these ratings yield a joint-standing score of 5.66, the square root of $(-4^2 + 4^2 = c^2)$. This joint-standing score (the distance of AC in Figure 1) represents the maximum distance from the ideal. Such a score represents extreme lopsidedness between the two items, or a complete absence of versatility with respect to an item pair. Joint-standing scores range between these two extremes. The Observed score (-2, 2), designated by point B in Figure 1, represents an intermediate value between the ideal rating and the extreme rating. The distance of line AB is equal to 2.83; this result is 50 % less versatile than a leader with an ideal score on the same pair of items.

Figure 2
Calculation of Joint-standing Scores



LVI-S sub-dimension sets are composed of four item pairs. Thus, each sub-dimension set produces four item-level joint-standing scores. These four item-level scores can be averaged to yield a joint-standing score for the sub-dimension set. Joint-standing scores are calculated for each of the three sets: *Takes Charge / Empowers Others*; *Declares / Listens*; and *Pushes / Supports*. Similarly, a joint standing score for the Forceful / Enabling duality is computed from the average of the three sub-dimension joint-standing values. Additional explanation of joint-standing score calculations is provided in Appendix X.

The Leadership Practices Inventory Student Version (SLPI)

The SLPI is one of the few leadership instruments designed for college students that has been validated within a collegiate context (Posner, 2005; Schwartz & Gimbel, 2000). Also, the SLPI is the only multi-rater leadership assessment instrument designed specifically for use with students (Leadership Challenge, 2008). The SLPI is a brief, 30-item instrument. Raters score their perceptions of a leader's behavior on five distinct "practices" of leadership: Modeling the Way, Inspiring a Shared Vision, Enabling Others to Act, Challenging the Process, and Encouraging the Heart (Posner, 2004). Each practice is measured with six items and each item is measured on a frequency scale that ranges from 1 (*rarely observed*) to 5 (*frequently observed*). The items are behavioral statements that characterize an action that represents an aspect of the five practices described above. A copy of the self-report version of these items is provided in Appendix A (items 25-54) and a copy of the observer-report version of these items is provided in Appendix B (items 25-54).

The instrument can be administered through paper and pencil or through a web-based administration process. Web-based instruments are scored automatically. Paper and pencil versions can be scored by hand, through the help of a student-focused workbook or the scores can be manually entered into a PC-based scoring program. The program will automatically score the instruments and a multi-page feedback report can be generated for the feedback recipient. In the present study, permission was granted to replicate the instrument in an electronic format.

Posner and Brodsky (1993) administered the SLPI to a group of RAs and their

residents. In the study 333 RAs participated and 1304 observers participated. On self ratings, factors exhibited a range of internal reliability coefficients (α): Modeling the Way = .69, Inspiring a Shared Vision = .81, Challenging the Process = .65, Enabling Others to Act = .69, Encouraging the Heart = .83 to a high of $\alpha = .83$. Observer ratings in this study were more stable, ranging from a low $\alpha = .81$ to $\alpha = .89$ across each of the five factors. In an unpublished thesis study (Pugh, 2000) the SLPI exhibited test-retest correlations exceeding $r = .51$ over a ten-week period (Posner, 2004). The instrument has demonstrated reliable psychometric properties across a variety of campus populations such as fraternity presidents, sorority presidents, resident assistants, orientation leaders, and general student samples (Posner, 2004).

Reputational Effectiveness Scale

Overall effectiveness ratings will serve as the dependent variable in this study. Tsui (1984) created a scale designed to measure reputational effectiveness. This scale measures the extent to which a manager has met the observer's performance expectations. Three items compose the scale: (1) "Overall, to what extent do you feel the manager is performing his job the way you would like it to be performed?" (2) "To what extent has he (focal manager) met your own expectations in his managerial roles and responsibilities?" and (3) "If you entirely had your way, to what extent would you change the manner in which he is doing his job?" Observers rate their manager using a 7-point Likert-type response scale (1 = "Not at all" to 7 = "Entirely"), and the third item is reverse-scored. Internal consistency estimates of reliability were reported for ratings of self, superiors, subordinates, and peers of $\alpha = .75, .84, .87,$ and $.86$ respectively.

Tsui's scale has been correlated with other effectiveness measures (Tsui, 1984). McCall and Segrist (1978) created a behaviorally focused questionnaire based on Mintzberg's (1973) classification of managerial roles. The roles measured were leader, liaison, entrepreneur, environment monitor, resource allocator, and spokesperson. Forty items were used to measure both importance and magnitude scales; the importance scale was used to measure role expectations. In addition to role expectations, Tsui (1984) measured a manager's relative reputation to others. A single-item construct was used (Tsui, 1984). On a 9-point scale, (1 = a great deal lower and 9 = a great deal higher), respondents indicated their response to the following question: "Relative to all other managers that you know in the company, what is your personal view of the reputation of this focal manager in terms of his overall effectiveness as a manager?" The reputational effectiveness scale was highly correlated with the six factors from McCall & Segrist (1978) role expectation questionnaire as well as the relative reputation measure (average $r = .56$). The 3-item scale also exhibited the highest reliability coefficients of the three (Tsui, 1984). A composite effectiveness variable (Eff_comp) was calculated from the cumulative score of all three effectiveness items.

Demographic Questionnaire

A brief demographic questionnaire (see Appendices B and C) was used to collect information on participants. Forms differed slightly for HDs (focal leaders) and RAs (observers). Forms inquired about the participant's years of experience working in a residence hall setting, type of setting (e.g., traditional residence hall), number of RAs

reporting to him/her (HD form), number of residents on hall (RA form) and demographic data including age, gender, racial affiliation, and institutional affiliation.

Figure 3
Instruments, Scales, and Subscales

| Instrument / Scale / Subscale | # Items |
|--|----------------|
| Leadership Versatility Index-Student (LVI-S) | 24 |
| <i>Forceful Dimension</i> | 12 |
| Takes Charge | 4 |
| Declares | 4 |
| Pushes | 4 |
| <i>Enabling Dimension</i> | 12 |
| Empowers Others | 4 |
| Listens | 4 |
| Supports | 4 |
| Student Leadership Practices Inventory (SLPI) (Kouzes & Posner, 2003) | 30 |
| <i>Models the Way</i> | 5 |
| <i>Inspires a Shared Vision</i> | 5 |
| <i>Challenges the Process</i> | 5 |
| <i>Enables Others to Act</i> | 5 |
| <i>Encourages the Heart</i> | 5 |
| Reputational Effectiveness Scale adapted of (Tsui, 1984) | 3 |
| <i>Reputational Effectiveness</i> | 3 |
| Demographic Questionnaire | - |
| <i>Demographic data (age, race, etc.)</i> | - |
| <i>Years experience</i> | - |
| <i>Type of residential setting</i> | - |
| <i>Number of direct reports (HD)</i> | - |
| <i>Number or residents on hall (RA)</i> | - |

Procedures

Modifying the LVI into the Leadership Versatility Index—Student (LVI-S)

The Leadership Versatility Index-Student (LVI-S) is a modification of Kaplan and Kaiser's (2006) Leadership Versatility Index, a corporate executive leadership development instrument based on the versatile leadership model (Kaplan & Kaiser, 2006). The LVI-S is a college student-focused, multi-rater measure of leadership versatility. In the LVI-S, multiple raters (observers) use an electronic form to provide quantitative feedback on their perceptions of their supervisor's behaviors. The LVI-S measures how well a supervisor balances forceful and enabling approaches to leadership. Supervisors with a high level of leadership versatility are adept at changing their approach to suit the situation and the personnel involved. Leaders with lower versatility exhibit a more lopsided approach. They tend to rely on a narrow range of options that limit their ability to adapt to the situation and the people involved.

The Leadership Versatility Index was modified into the Leadership Versatility Index for Students through a multiple-wave development process. The first involved a review of student leadership literature, literature on the training and development of paraprofessional and professional leaders in housing and residence life communities, as well as a review of the versatile leader model and Kaplan and Kaiser's (2006) Leadership Versatility Index. This literature is summarized in Chapter II.

Establishing the Scope of the LVI-S

An initial review of the Leadership Versatility Index determined that the Forceful / Enabling duality, which measures aspects of interpersonal influence, was appropriate

for a campus-based leadership instrument. Guidelines for training and developing paraprofessional supervisors highlighted the importance of Forceful and Enabling leadership skills (Upcraft & Pilato, 1982; Winston & Fitch, 1993). Literature indicated all of the Forceful / Enabling sub-dimensions were relevant to the housing and residence-life context. Finally, the review noted the importance of establishing a synergistic balance between Forceful and Enabling leadership behaviors (Winston, & Fitch, 1993; Winston Ullum, & Werring, 1984). The Strategic / Operational duality was excluded because these leadership behaviors were not relevant to entry-level leadership contexts. Paraprofessionals do not set direction or outline operational processes for departments in housing and residence life—they apply structure according to guidelines issued from the departments.

Item Modifications for the LVI-S

The second wave of development addressed item modification. Items were modified according to a four-step process. First, items from the LVI were reviewed and edited to make them more suitable for a student-centered context. This editing process produced an initial draft of LVI-S items. These items are presented in Appendix F. Next, a focus group comprised of student affairs and housing and residence life experts was convened. These experts reviewed items and provided feedback on their relevance to paraprofessional leadership. They also considered whether item language was suitable for a campus setting. This expert focus group consisted of four members: two professional housing and residence life administrators (one Caucasian male and one African-American female, both with Masters' degrees); a female faculty member (Ph.D., Caucasian) who

researches in the field of higher education and teaches masters- and doctoral-level courses in higher education; and a male, African-American student affairs administrator who obtained a Ph.D. in higher education, has over 20 years of professional experience in student affairs administration, five of these in housing and residence life. Feedback was provided in written and verbal forms.

Most items were deemed acceptable, but several were substantially modified. These modifications were due to interpretive differences between business and educational contexts. Item 9 (Fp_9) was originally worded, “*Pushes others hard.*” In business settings, bosses have a higher degree of coercive influence than do paraprofessionals in a residence life setting, thus they can “push others hard” to drive business objectives. Experts believed “*challenges people to do their best—Pushes hard for high performance*” was more appropriate for the campus context. Item 20 (E1_20) was problematic as well. The original wording, “*Makes it easy for people to challenge his/her thinking*” was supposed to reference a leader who was open to debate or provided adequate space for others to challenge their thinking. Several experts interpreted the item as meaning the leader appeared incompetent (thus making it easy for someone to challenge his/her thinking). Since this interpretation differed significantly from the original intent, the item was changed to “*Makes it safe to challenge or critique his/her thinking—welcomes debate.*” Three other items (Ee_13, Ee_14, and Ee_22) were modified in subtle ways. Each item modification may be traced in Appendix E.

After processing the expert feedback, a second draft of items was generated and a Resident Advisor focus group was convened. This group consisted of three participants,

all female. One member was African-American and the other two were Caucasian. Male RAs were recruited and expressed interest in participating but were unavailable for the date of the meeting. This lack of male participation may be considered a limitation of this focus group procedure. The second draft of items was presented to the RA focus group. As in the previous expert focus group, feedback was collected in written and verbal form. The RA focus group represented the population who would be completing the instrument, their feedback helped to validate whether the intended population interpreted items as intended. Most modifications at this stage involved minor adjustments to wording, but one change warrants further discussion. Item 4 (Ftc_4) was initially worded, “*Steps in—gets personally involved when problems arise.*” This item was intended to represent a HD who was willing to put immediate responsibilities on hold to personally address issues occurring on the RA’s hall. Participants in the RA focus group, however, interpreted the item differently. These RAs received extensive training on respecting residents’ boundaries and not becoming personally involved in their problems. Perhaps as a result of this training, the RAs interpreted getting “personally involved” to mean the leader was crossing interpersonal boundaries. To alleviate this misunderstanding, the item was modified to state, “*Steps in—gets actively involved when problems arise.*” Four other items were modified (Ftc_3, Fd_6, Fd_8, and El_17). As noted before, Appendix D charts the evolution of each item modification.

A third draft of items was compiled after reviewing and processing the RA focus group feedback. This draft was submitted to a panel of experts familiar with the Too Little/Too Much rating scale and the duality-based structure of the LVI. Each of these

experts possessed graduate degrees in Industrial/Organizational Psychology; one held a Master's degree and the other two held a Ph.D. degree. Two were psychometric specialists familiar with design and application of the LVI, the third had more than 30 years experience in executive development and feedback and has practiced extensively with the LVI. These experts were asked to assess the third draft of items and report whether they believed the proposed item modifications could threaten the dimensional integrity of the Forceful / Enabling duality.

The dimensional integrity of the Forceful / Enabling duality is maintained by ensuring that forceful items do not contain enabling language and vice-versa. The evolution of item Ftc_3 exemplifies a potential threat to dimensional integrity. The initial item read, "*Sets clear expectations—tells people what to do.*" Later in the review process, RAs preferred this item to read, "*Delegates clearly—tells people what to do.*" "Sets clear expectations" and "delegates clearly" were assumed to be nearly synonymous phrases. Expert reviewers cautioned that "delegation" was an enabling leadership behavior. A leader delegates responsibility to someone else, thus enabling that person to operate as a leader. The item was revised to read, "*Gives clear direction, tells people what to do.*" "Gives clear direction" better represented the forceful dimension because it implies the leader is instructing with authority, rather than delegating authority to someone else. This expert panel suggested modifications to six additional items (Fp_12, Ee_13, El_20, Es_21, Es_22, and Es_24). The evolution of these items can be reviewed in Appendix C.

A fourth draft of items was generated subsequent to the LVI panel's expert review. This fourth draft of items was used in the final step of the LVI to LVI-S

modification process. The field test instrument consisted of 24 Forceful / Enabling items and two sets of effectiveness items: 10 behavioral effectiveness items used by Posner & Brodsky (1993), and three global items that measured reputational effectiveness (Tsui, 1984).

A convenience sample of two HDs and their respective RAs (four per hall director) were recruited ($n = 10$). The field test helped to identify troublesome items, hone instrument instructions, and to clarify test administration guidelines. Upon completion of the assessment, the administrator provided a brief survey of open-ended questions to gather participants' reaction to the instrument (e.g., "Were the instructions clear?", "What items were unclear to you?", "Was the response format easy for you to understand?", and "How long did it take you to complete the instrument?"). Generally, participants believed the questions were clear. Two commented they would have preferred a more conventional response scale because they found the Too Little Too Much scale confusing at first. The pilot instrument provided instructions that addressed the reasons a rater might choose the "too much" rating or the "too little" rating, but did not explicitly address the "right amount." One respondent requested additional instructions that clarified what "the right amount" should represent. Feedback from the pilot study was incorporated into the final version of the instrument used in the main study (Appendix A). The ten effectiveness items from Posner and Brodsky (1993) were deemed confusing and not well suited for this research application. These items were eliminated in favor of the three-item scale derived from Tsui (1984).

Development of LVI-S Assessment Platform

Based on experience gained from administering the pilot study, it was determined the final instrument should be electronically administered. Managing a pilot sample of 10 participants was feasible, but scaling the paper and pencil process to include 8 institutions, more than 50 HDs, and approximately 250 RAs was not.

Commercial LVI assessment platform. The LVI utilizes an electronic assessment platform customized specifically for the instrument and the Too Little Too Much (TLTM) response scale. These customizations are necessary to control for potential threats to validity. The TLTM scale has its most favorable rating, “*the right amount*” located in the middle of the scale. To indicate behaviors are used “*too little*,” respondents should mark values to the left of center; to indicate behaviors used “*too much*,” respondents should mark values to the right of center. Kaiser (2009, personal communication) warned respondents sometimes overlook or forget these criteria. They mark values at the end of the scale (in the +3 or +4 range), assuming these to be the most favorable response options. This type of error can present a significant threat to validity. The commercial platform uses passive and active measures to control this threat. Graphic and textual reminders, visible throughout the administration of the assessment, serve as passive control measures. An active control measure is triggered when a rater provides several consecutive responses in the “*much too much*” range. Participants receive a message confirming whether or not they meant to indicate over-use of the behavior. This message reiterates instructions for the TLTM response format and advises the participant to please revise ratings if the scale had been misinterpreted.

LVI-S administration platform. The LVI-S was not administered on the commercial LVI platform. The TLTM scale precluded the use of online survey software (e.g., SurveyMonkey) because these products could not incorporate the passive and active control measures necessary for administering the LVI-S. A custom platform was designed. This platform consisted of a website that hosted web-based informed consent documents, HD and RA survey forms (with the appropriate passive and active measures for controlling accidental over-rating), an enrollment form for the iPod Shuffle drawing, three separate, password-protected databases (one for HD response data, one for RA response data, and one to record entries for the drawing). A Secure Socket Layer (SSL) certificate protected the data as it was transmitted from the forms to the databases; SSL technology is used to encrypt and protect sensitive financial information and financial transactions such as credit-card purchases. To protect respondents' anonymity, an identity obfuscation algorithm was programmed into the HD and RA response forms. A complex mathematical process converted participants' identifying information into a unique code consisting of approximately 30 characters, which used upper and lowercase letters as well as numbers. Participant anonymity was maintained, but the codes permitted data to be organized into the HD-RA rater groups necessary for analysis. The web-based platform was complemented by a contact management spreadsheet maintained in Excel 2004, a series of Entourage 2004-based e-mail templates used for participant recruitment and follow-up. The platform was beta-tested to ensure all pages, page links, forms, control measures, databases, and security measures operated as intended. Once all platform elements passed the beta test, it was deemed ready to administer the main study.

Administration of the Study

A convenience sample of HDs and RAs was recruited through the housing and residence life departments at 7 institutions of higher education located in the Southeastern United States. Participating schools included four public universities, two private colleges, and an HBCU. Initial contact was made with Directors or Assistant Directors in the housing and residence life departments of participating institutions. An overview of the study was provided, and departments were informed that upon completion of the study, an on-site mini-colloquium would review the results of the study. Once departments agreed to participate, they were asked to sign a letter of support for the study. Letters of support have been recorded in Appendix S. After submission of the support letter, departments were asked to complete a contact information template. These templates were assembled to create a master contact list used to monitor completion rates and issue follow-up reminders.

An initial recruitment message was sent to the housing directors and the directors forwarded this message to their Hall Directors through their inter-departmental distribution list. Participants were informed of the voluntary nature of the study, the potential benefits and costs of participation, and were provided a link to the informed consent page on the LVI-S website. Once the link was clicked, the HD was presented with an electronic Hall Director informed consent document. They were encouraged to print a copy of the document for their records. Hall Director participants were informed that they could provide consent by clicking the study link at the bottom of the informed consent document.

Hall Directors completed the self-form of the instrument and were instructed to the submit button. Once the data was submitted, the HD received a pre-written RA recruitment message; they were asked to cut and paste this message into an e-mail addressed to the RAs they supervised. This e-mail informed RAs of their supervisor's participation in the study and that voluntary feedback on their supervisor's leadership would be appreciated. At the same time HDs received the RA recruitment message, they also were provided a link that would take them to the registration form for the iPod Shuffle drawing. Hall Directors then received a thank you message for participating.

The RA recruitment message sent by the HD referenced the benefits and costs of participation and provided a link to the RA informed consent document. Clicking the link directed RAs to the consent form. They were encouraged to print a copy of the document for their records and were informed that they could provide consent by clicking the study link at the bottom of the consent document. Once this link was activated, RAs were directed to the RA observer form that contained all the instruments included in the study. Once this form was completed and submitted, the RA was directed to the registration form for the iPod Shuffle drawing and was thanked for participating in the study.

Follow-up messages were issued to non-responders on a weekly basis after the beginning of data collection. The data collection period for most schools in the study was six weeks, however one institution was approved to join the study during the middle of the data collection process and this institution's data collection period was two weeks shorter than those for the other institutions.

Data Analysis

The research questions have been designed to progressively explore the psychometric soundness of the LVI-S. Initially, descriptive statistics were analyzed through the use of the Statistical Package for the Social Sciences (PASW Statistics; Version 18.0 for Apple Macintosh). Reliability coefficients for each factor were calculated. The LVI-S structure was examined through exploratory factor analysis. Regressions were performed to determine which LVI-S sub-dimensions predicted HD effectiveness. Correlations determined the empirical relationship between the LVI-S sub-dimensions and the SLPI scales. A cross-tab analysis organized responders into agreement categories. A series of ANOVA analyses determined whether significant differences in effectiveness ratings could be discerned across agreement categories. Post hoc analyses helped determine the direction of significant differences. A final series of regression analyses compared the predictive power of an additive regression model that did not take joint-standing scores into account versus a regression model that accounted for joint-standing scores.

Table 1
Research Questions

| | | |
|--|--|------------------------------|
| Research Question 1: Does the LVI-S verify that there are two leadership factors: forceful leadership and enabling leadership? | | |
| Hypothesis | Variables | Analysis |
| <i>Hypothesis 1:</i> The LVI-S will present a clearly identifiable two-factor structure. These two factors will have an inverse relationship to one another. | Sub-dimensions of forceful and enabling leadership | EFA, Correlation Analysis |
| <i>Hypothesis 1b:</i> The LVI-S will provide adequate empirical evidence for a duality-based hierarchical structural model composed of 1 dimension and 3 sub-dimensions for each side of the F/E duality. | Sub-dimensions of forceful and enabling leadership | EFA, Correlation Analysis |
| Research Question 2: What proportion of the HD reputational effectiveness can be explained by the six LVI-S sub-dimensions? Are any of the six sub-dimensions statistically significant predictors of HD reputational effectiveness as reported by RAs? | | |
| Hypothesis | Variables | Analysis |
| <i>Hypotheses 2a:</i> Scores associated with the items for each sub-dimension (i.e., HD sub-dimension scores, as reported by RAs), will be statistically significant predictors of leader effectiveness ratings. | HD F/E sub-dimension scores (<i>Takes Charge, Declares, Pushes, Empowers, Listens, and Supports</i>) as reported by RAs HD Reputational Effectiveness Scores as Reported by RAs | Regression Analysis; t-Test |
| <i>Hypothesis 2b:</i> HD self-reported sub-dimension scores will not be statistically significant predictors of leader effectiveness ratings. | HD self-reported F/E sub-dimension scores (<i>Takes Charge, Declares, Pushes, Empowers, Listens, and Supports</i>) HD self-reported Reputational Effectiveness Scores | Regression Analysis; t-Test, |
| Research Question 3: To what extent do scores on the LVI-S and the SLPI correlate, providing evidence of convergent construct validity of the LVI-S? | | |
| <i>Hypotheses 3a:</i> LVI-S sub-dimensions of Enabling leadership significantly correlate with Enabling Others to Act, Inspiring a Shared Vision, and Encouraging the Heart on the SLPI. | SLPI Variables: Encouraging the Heart, Enabling Others to Act, and Inspiring a Shared Vision LVI-S Sub-dimensions of Enabling leadership (<i>Empowers, Listens, and Supports</i>) | Correlation Analysis |

| | | |
|--|--|---|
| <i>Hypothesis 3b:</i> LVI-S sub-dimensions of Forceful leadership will significantly correlate with the Challenging the Process variable on the SLPI. | LPI Variable: Challenging the Process LVI-S Sub-dimensions of Forceful leadership (<i>Takes Charge, Declares, Pushes</i>) | Correlation Analysis |
| Research Question 4: Do most HDs underestimate or overestimate their F/E dimension scores in comparison to RAs ratings of the HD's F/E dimension scores? | | |
| Hypothesis | Variables | Analysis |
| <i>Hypothesis 4:</i> Most HDs will overestimate their leadership scores on the six F/E sub-dimensions (rate themselves closer to the right amount) than the Resident Advisors' ratings of HD. | HD F/E dimension scores HD self-reported F/E sub-dimension scores | Mean Difference Score Comparison & Crosstab Analysis |
| Research Question 5: Is there a statistically significant mean difference in effectiveness ratings between HDs who overestimate their F/E dimension scores versus HDs who underestimate their F/E dimension scores? | | |
| Hypothesis | Variables | Analysis |
| <i>Hypothesis 5a:</i> There will be a statistically significant mean difference in effectiveness ratings between HDs who overestimate their F/E dimension scores versus those who underestimate their F/E dimension scores. <i>Hypotheses 5b:</i> HDs who overestimate their F/E dimensions scores will be rated less effective by RAs. | Over-rater and Under-rater HD grouping variable (as determined through difference between average of RA reported HD dimension scores and self-reported HD dimension scores) HD composite effectiveness ratings (Eff_comp) as reported by RA | ANOVA for agreement groups and effectiveness scores; Scatterplot comparing HD & RA_average dimension scores and effectiveness ratings; Correlation analysis |
| Research Question 6: Which regression model better predicts reputational effectiveness scores: (a) an additive model where interaction between complementary F/E sub-dimensions is not accounted for, or (b) a model that incorporates the focal leader's joint standing scores (versatility) on complementary F/E sub-dimensions? | | |
| <i>Hypothesis 6:</i> Model 2 will explain a greater percent of the variance in reputational effectiveness; model 2 will more effectively predict reputational effectiveness scores. | HD F/E sub-dimension scores (<i>Takes Charge, Declares, Pushes, Empowers, Listens, and Supports</i>) as reported by RAs HD Reputational Effectiveness Scores as Reported by RAs | Regression analysis |

CHAPTER IV

RESULTS

In this chapter, the results of the study are presented using descriptive and inferential statistics. First, the characteristics of the sample are described. Second, the psychometric information is reported for each instrument. Finally, the analyses to test for each research hypothesis are described and results provided.

Sample Characteristics

Hall directors ($n = 57$) and resident advisors ($n = 262$) participated in the study. Participants were employed by one of seven colleges and universities in the southeastern United States. Institutions varied in size and in public/private status, including two religiously affiliated private institutions with enrollments that did not exceed 5000 students, an HBCU with enrollment between 5000-10,000 students, two mid-sized public universities with enrollments between 10,000-20,000 students, and two large institutions with enrollments that exceeded 20,000 students. These institutions employed a variety of residence types, including traditional residence halls, learning communities, apartments, and suite-style accommodations. Residents included undergraduate and graduate students.

Hall Directors

Hall Directors (HDs) ranged in age from 23 to 64 ($m = 28.23$, $sd = 5.93$). The sample included 27 men (47.4%) and 30 women (52.6%). Forty-five respondents

(81.82%) were from public universities, 4 (7.27%) were affiliated with an HBCU, and 6 (10.91%) came from private colleges. HDs' experience ranged from 0 to 11 years ($m = 2.84$, $sd = 1.97$). The number of RAs supervised by an HD ranged from 3 – 20 ($m = 10.07$, $sd = 3.63$). The average number of raters-per HD was 4.92 and the range of feedback providers spanned from 1 to 13 ($sd = 2.9$). A summary of demographic and categorical characteristics of the HD sample has been provided in Tables 2 - 7 below.

Table 2
Hall Director Sample Characteristics

| Continuous Variables | N | Minimum | Maximum | Mean | <i>sd</i> |
|--|----|---------|---------|-------|-----------|
| Age (in Years) | 57 | 23 | 64 | 28.23 | 5.93 |
| Number of years with RA Experience | 38 | 1 | 3 | 2.42 | .72 |
| Years experience as a Hall Director | 57 | 0 | 11 | 2.84 | 2.0 |
| Number of RAs supervised | 55 | 3 | 20 | 10.07 | 3.63 |
| Number of observers providing feedback | 53 | 1 | 13 | 4.92 | 2.90 |

Table 3
Distribution of HD Sample Across Institutional Type/Residence Type

| Institution/Residence Type | Size (in student population) | | | | Total | % Total | Cum. % |
|--------------------------------|------------------------------|----------|-----------|-----------|-----------|-----------------|-----------------|
| | 1K-5K | 5K-10K | 10K-20K | >20K | | | |
| Public University | | | | | | Public Univ. | Public Univ. |
| Traditional Hall | - | - | 6 | 13 | 19 | 34.55 | 34.55% |
| Learning Comm. Combination | - | - | 2 | - | 2 | 3.64 | 38.18% |
| Trad/LC | - | - | 6 | 13 | 19 | 34.55 | 72.73% |
| Other | - | - | 2 | 3 | 5 | 9.09 | 81.82% |
| <i>Public University Total</i> | - | - | 16 | 29 | 45 | 81.82 | 81.82% |
| HBCU | | | | | | HBCU | HBCU |
| Traditional Hall | - | 3 | - | - | 3 | 5.45 | 87.27% |
| Learning Comm. Combination | - | - | - | - | - | - | - |
| Trad/LC | - | 1 | - | - | 1 | 1.82 | 89.09% |
| Other | - | - | - | - | - | - | - |
| <i>HBCU Total</i> | - | 4 | - | - | 4 | 7.27 | 89.09% |
| Private College | | | | | | Private College | Private College |
| Traditional Hall | 6 | - | - | - | 6 | 10.91 | 100% |
| Learning Comm. Combination | - | - | - | - | - | - | - |
| Trad/LC | - | - | - | - | - | - | - |
| Other | - | - | - | - | - | - | - |
| <i>Private College Total</i> | 6 | - | - | - | 6 | 10.91 | 100% |
| Total | 6 | 4 | 16 | 29 | 55 | 100% | 100% |

Table 4
Number of RAs Supervised Versus Number of RA Participants per Institution Type

| Institution Type | # RAs Supervised | | | | # RAs Providing Observations | | | |
|----------------------|------------------|--------------|-------------|-------------|------------------------------|-------------|-------------|-------------|
| | Mean | Max. | Min. | SD | Mean | Max. | Min. | SD |
| Large Public | 11.70 | 17.50 | 8 | 2.60 | 5.56 | 13 | 1 | 3.35 |
| Medium Public | 8.79 | 12 | 5 | 2.31 | 3.6 | 7 | 1 | 1.88 |
| Small Private | 6.63 | 9 | 4 | 3.08 | 3.6 | 6 | 1 | 2.07 |
| HBCU | 8.75 | 14 | 4 | 4.11 | 3.67 | 5 | 2 | 1.71 |
| Total Average | 8.97 | 13.13 | 5.38 | 3.03 | 4.11 | 7.75 | 1.25 | 2.95 |

Table 5
 HD Participation Rate by Institutional Type

| Institution Type | # HDs Sampled | # HDs Responded | Participation Rate |
|-------------------|---------------|-----------------|--------------------|
| Public University | 58 | 45 | 77.6% |
| HBCU | 7 | 4 | 57.1% |
| Private College | 12 | 7 | 58.3% |
| Total | 77 | 56 | 72.7% |

Table 6
 Race/Ethnicity of Hall Directors by Institution and Residence Hall Type
Institution Type

| Ethnicity | <i>Institution Type</i> | | | Total | % of Total | Cumulative % of Total |
|---------------------------|-------------------------|----------|-----------------|-----------|------------|-----------------------|
| | Public University | HBCU | Private College | | | |
| White | 33 | - | 3 | 36 | 63.2 | 63.20 |
| Black or African-American | 9 | 4 | 2 | 15 | 26.30 | 89.50 |
| Bi-Multiracial | 2 | - | - | 2 | 3.5 | 93 |
| Native American | 1 | - | - | 1 | 1.80 | 94.70 |
| Asian | 1 | - | - | 1 | 1.80 | 96.50 |
| Pacific Islander | 1 | - | - | 1 | 1.80 | 98.20 |
| Other | - | - | 1 | 1 | 1.80 | 100 |
| Total | 47 | 4 | 6 | 57 | 100 | 100 |

Table 7
 Gender of Hall Directors by Institution and Residence Hall Type

| | Public University | HBCU | Private College | Total | % of Total |
|--------------|-------------------|----------|-----------------|-----------|------------|
| Male | 24 | 1 | 2 | 27 | 47.40 |
| Female | 23 | 3 | 4 | 30 | 52.60 |
| Total | 47 | 4 | 6 | 57 | 100 |

Resident Advisors

Resident Advisors (RAs) ranged in age from 20 to 30 ($m = 21.24$, $sd = 1.24$). The range of RA experience spanned from 1 to 4 years ($m = 1.53$, $sd = .71$). Two hundred and thirty-two (88.89%) RA participants were enrolled in public universities, 11 (4.2%) were enrolled in a HBCU, and 18 (6.9%) were enrolled in private colleges. A summary of demographic and categorical characteristics of the RA sample has been provided in Tables 8 - 12 below.

Table 8
RA Sample Characteristics

| Continuous Variables | N | Minimum | Maximum | Mean | <i>sd</i> |
|----------------------|-----|---------|---------|-------|-----------|
| Age (in Years) | 248 | 20 | 30 | 21.24 | 1.24 |
| Yrs. experience | 260 | 1 | 4 | 1.53 | .71 |
| # Residents | 257 | 10 | 1200 | 86.56 | 132.41 |

Table 9
Distribution of RA Sample Across Institutional Type/Residence Type

| Institution/Residence Type | Size (in student population) | | | | Total | %Total | Cum. % Total |
|--------------------------------|------------------------------|-----------|-----------|------------|------------|-----------------|-----------------|
| | 1K-5K | 5K-10K | 10K-20K | >20K | | | |
| Public University | | | | | | Public Univ. | Public Univ. |
| Traditional Hall | - | - | 31 | 117 | 148 | 56.7% | 56.7% |
| Learning Comm. | - | - | 8 | 37 | 45 | 17.24% | 73.95% |
| Other | - | - | 15 | 24 | 39 | 14.94% | 88.89% |
| <i>Public University Total</i> | - | - | 54 | 178 | 232 | 88.89% | 88.89% |
| HBCU | | | | | | HBCU | HBCU |
| Traditional Hall | - | 8 | - | - | 8 | 3.07% | 91.95% |
| Learning Comm. | - | 3 | - | - | 3 | 1.15% | 93.1% |
| Other | - | - | - | - | - | - | - |
| <i>HBCU Total</i> | - | 11 | - | - | 11 | 4.2% | 93.1% |
| Private College | | | | | | Private College | Private College |
| Traditional Hall | 18 | - | - | - | 18 | 6.9% | 100% |
| Learning Comm. | - | - | - | - | - | - | - |
| Other | - | - | - | - | - | - | - |
| <i>Private College Total</i> | 18 | - | - | - | 18 | 6.9% | 100% |
| Total | 18 | 11 | 54 | 178 | 261 | 100% | 100% |

Table 10
RA Participation Rate by Institutional Type

| Institution Type | # RAs Sampled | # RAs Responded | Participation Rate |
|-------------------|---------------|-----------------|--------------------|
| Public University | 459 | 232 | 50.5% |
| HBCU | 42 | 11 | 26.1% |
| Private College | 35 | 18 | 51.4% |
| Total | 536 | 261 | 48.7% |

Table 11
Race/Ethnicity of RAs by Institution
Institution Type

| Ethnicity | <i>Institution Type</i> | | | Total | % of Total | Cum. % of Total |
|---------------------------|-------------------------|-----------|-----------------|------------|------------|-----------------|
| | Public University | HBCU | Private College | | | |
| White | 156 | - | 12 | 168 | 64.1 | 64.1 |
| Black or African-American | 44 | 11 | 1 | 56 | 21.4 | 85.5 |
| Bi-Multiracial | 9 | - | 1 | 10 | 3.8 | 89.3 |
| Native American | 2 | - | - | 2 | .8 | 90.1 |
| Asian | 14 | - | - | 14 | 5.3 | 95.4 |
| Pacific Islander | - | - | - | - | 0.0 | 95.4 |
| Not Identified | 1 | 1 | 2 | 4 | 1.5 | 96.9 |
| Latino/a | 3 | - | - | 3 | 1.1 | 98.1 |
| Other | 3 | - | 2 | 5 | 1.9 | 100 |
| Total | 232 | 12 | 18 | 262 | 100 | 100 |

Table 12
Gender of Resident Advisors by Institution
Institution Type

| Gender | <i>Institution Type</i> | | | Total | % of Total | Cum. % of Total |
|----------------|-------------------------|-----------|-----------------|------------|------------|-----------------|
| | Public University | HBCU | Private College | | | |
| Male | 96 | 4 | 7 | 107 | 40.8 | 40.8 |
| Female | 135 | 7 | 10 | 152 | 58 | 98.8 |
| Not Identified | 1 | 1 | 1 | 3 | 1.1 | 100 |
| Total | 232 | 12 | 18 | 262 | 100 | 100 |

Instrument Psychometrics

HD and RA survey responses were recorded via similar, but separate electronic forms and data for each group was stored in separate databases. Psychometric properties of the LVI-S were calculated using the observer database because observer ratings have been found more reliable than self-ratings (Conway & Huffcutt, 1997). Each observer

record contained a Hall Director ID Code in addition to the Resident Advisor ID Code. Records were hierarchically sorted by Hall Director ID code and then Resident Advisor ID code, thus producing the leader-observer rater groups used for psychometric analysis of the instrument.

Because all the data were collected using a single form, some control for common method bias was warranted. Common method bias could occur when raters have not discriminated between differing constructs within an instrument. Excessive amounts of shared covariance between unrelated variables could make it difficult to differentiate between measurement artifacts and the phenomena under investigation (Avolio & Bass, 1991; Malhotra, Kim, & Patil, 2006). The marker variable technique helped control for common-method bias (Malhotra, Kim, & Patil, 2006). In the marker variable technique, a variable believed to be theoretically unrelated to an outcome variable was selected and a correlation analysis was performed. Ideally, the analysis should indicate very little or no relationship between the variables. If a correlation did exist, this result could indicate a problematic level of common method variance. In the current study, the rater confidence variable was correlated with the effectiveness composite variable ($r = .02$). This result provided a measure of confidence that the data collected by the LVI-S forms were not unduly influenced by common method bias.

Psychometric testing is an inexact science and some error is expected. Reliability is an estimate of this characteristic. Following recommendations provided by LeBreton, Burgess, Kaiser, Atchley, and James (2003), reliability was examined from multiple

perspectives: internal consistency (Chronbach's alpha) and the level of consistency and consensus among raters (Intraclass Correlation Coefficient).

Chronbach's alpha was used to calculate the degree of internal consistency of each scale. According to Nunnally (1978), alphas greater than .70 are acceptable in psychological assessments but alphas of .80 or higher are preferred. The 12-item Forceful dimension produced an acceptable internal consistency value ($\alpha = .83$). Alphas for the three sub-dimensions comprising the Forceful dimension, however, were lower. An item-level examination indicates that alpha values do not appear to be compromised by a single item. Appendix Y provides inter-correlation matrices from the items within each sub-dimension. *Takes Charge* exhibits low inter-item correlations across the board (mean $\Phi = .30$, minimum $\Phi = .13$, maximum $\Phi = .52$; $\alpha = .64$). The *Declares* sub-dimension is similarly afflicted, though not to the same extent (mean $\Phi = .33$, minimum $\Phi = .17$, maximum $\Phi = .48$; $\alpha = .68$). The *Pushes* sub-dimension provides a stark contrast to the other two (mean $\Phi = .44$, minimum $\Phi = .30$, maximum $\Phi = .63$; $\alpha = .76$). The higher inter-item correlation contributed greater internal consistency within the sub-dimension.

The Enabling dimension produced an alpha coefficient that exceeds the preferred standard ($\alpha = .83$). As expected, alpha scores for the sub-dimensions were lower (Table 14). As before, no single item appeared to dramatically impact alpha scores for the Enabling sub-dimensions. The weaker alpha in the *Supports* sub-dimension may be traced to lower inter-correlation within the sub-dimension (mean $\Phi = .32$, minimum $\Phi = .19$, maximum $\Phi = .43$; $\alpha = .69$).

LeBreton et al. (2003) recommended using inter-rater reliability and inter-rater agreement to measure the level of consistency and consensus at the rater-group level. Inter-rater reliability was calculated using the one-way random effects Intraclass Correlation Coefficient (ICC; Shrout & Fliese, 1976). The one-way random effects ICC exists in two forms: ICC(1,1) and ICC(1, k) (LeBreton et al, 2003). The ICC(1,1) version provides insight into the reliability of individual raters. It indicates whether individual raters provide information about their targets that is reliable and consistent with other raters—whether ratings exhibit consistent patterns of highs and lows. The second version, ICC(1, k) provides insight into the consistency of mean ratings for a group of k raters.

LeBreton, et al., (2003) noted that $ICC(k)$ values will increase as k increases, however a balance must be negotiated between statistical ideals and sampling availability. A higher observers-per-leader ratio is preferable (greater k), but setting the number too high can reduce the number of groups available for analysis. The mean level of observers-per-rater in this study was 5. Using $k = 5$ as a starting point, rater-groups with 5 or more raters were selected for analysis. This selection criterion yielded 24 rater groups (24, $n = 160$). After increasing k to six and selecting for rater groups with six or more raters, the number of groups available was reduced by nearly a third and n was reduced by 25% (16, $n = 120$). Consequently, $k = 5$ was deemed to provide the highest $ICC(k)$, while optimizing the number of rater groups and participants available for analysis.

When comparing the $ICC(1,5)$ sub-sample to the overall sample, the ICC subset differed in some respects, but percentages of representation were similar in most

categories. In terms of gender, the original sample of RAs was biased toward females (58% to 41%). The *ICC* sub-sample was also biased but to a smaller degree (55% to 44%). At 18% and 67% respectively, African-Americans and Whites were within 3 percentage points of the original sample's participation rates (African-Americans were slightly lower, Whites were slightly higher). All other racial / ethnic groups accounted for approximately 14% of the *ICC* subset, within one percentage point of the overall sample. The largest differences between the *ICC* subset and the overall sample were reflected in the percentages of institutional representation. As in the overall study, large public institutions were represented to the greatest extent. In the *ICC* subset, large institution representation increased significantly (original sample 68%, *ICC* sample 82.5%). All other institutions were represented to a lower degree (Medium Public, 23% vs. 14.4%; Small Private, 7.8% vs. 3.1%; and HBCU, 4% vs. 0%). This shift in institutional participation is understandable because large universities had a 30% greater RA to HD ratio, providing more potential raters per leader (see Table 4). This advantage transferred into a greater number of observations received per leader, thus influencing the large institution bias in the *ICC* subset. Demographic characteristics and descriptive statistics for the *ICC* sub-sample are summarized in (Table 13). Table 14 provides a crosstab analysis of RA Year (sophomore, junior, senior) by RA Experience (years of experience). Although this data has not been compared to the overall dataset, it does provide useful insight into the experience level of participants contributing to the inter-rater reliability study.

Table 13
ICC Analyses Sub-sample: Demographic Characteristics and Descriptive Statistics

| | | Categorical Variables | | | | | | | |
|---------------------|----------|------------------------------|-------|---------------|-------|---------------|-------|-------|-------|
| Gender | | Female | | Male | | Null | | Total | |
| | <i>n</i> | 88 | | 71 | | 1 | | 160 | |
| | % | 55 | | 44 | | .6 | | 100 | |
| | | | | | | | | | |
| Race / Ethnicity | | African- American | Asian | Latino/a | Mixed | Null | Other | White | Total |
| | <i>n</i> | 30 | 9 | 2 | 7 | 2 | 2 | 108 | 160 |
| | % | 18.8 | 5.6 | 1.3 | 4.4 | 1.3 | 1.3 | 67.5 | 100 |
| | | | | | | | | | |
| Institution | | Large Public | | Medium Public | | Small Private | | Total | |
| | <i>n</i> | 132 | | 23 | | 5 | | 160 | |
| | % | 82.5 | | 14.4 | | 3.1 | | 100 | |

Table 14
ICC Sub-sample Crosstab Analysis: RA Year by RA Experience (in years)
Yrs. RA Experience

| RA Year | 1 yr. | 2 yrs. | 3 yrs. | 4 yrs. | Total |
|----------------|-----------|-----------|----------|----------|------------|
| Junior | 18 | 35 | 0 | 0 | 53 |
| Senior | 10 | 15 | 1 | 0 | 44 |
| Sophomore | 62 | 0 | 0 | 1 | 63 |
| Total | 90 | 50 | 1 | 1 | 160 |

In multi-rater instrumentation *ICC*(1) values typically range from .00 to .50 with an average of .12 (James, 1982). When values meet or exceed .05, they are considered high enough to aggregate (Bliese, 2000). No multi-rater data on *ICC* scores for college student populations was available, but Conway and Huffcut (1997) reported *ICC*(1) values that ranged from .2 to .45 in professional multi-rater studies. In this study, *ICC*(1) values on the LVI-S ranged from .16 to .37, values on the SLPI ranged from .35 to .58, and the *ICC*(1) value for Tsui’s reputational effectiveness scale was .64. Conway and

Huffcut (1997) reported $ICC(k)$ values in studies on professional managers ranging from .5 for $k = 3$ raters to .6 for $k = 5$ raters. The preferred cutoff values for $ICC(k)$ is .70, though values between .5 and .7 are typically considered acceptable (Kaiser & Kaplan, 2006). ICC scores for the LVI-S dimensions were (Forceful $k=5$, .60; Enabling $k=5$, .61) both below the preferred value but within the acceptable range. In four out of 6 sub-dimension cases, $ICC(5)$ scores were below .5. Table 15 presents the $ICC(1)$ and $ICC(5)$ results for the dimensions and sub-dimensions of the LVI-S, the scales of the SLPI, and Tsui's reputational effectiveness scale.

Table 15
Internal Consistency and Inter-rater Reliability

| <u>INSTRUMENT / Scale / Sub-scale</u> | <u># items</u> | <u>α</u> | <u>$ICC(1)$</u> | <u>$ICC(5)$</u> |
|---|----------------|----------------------------|----------------------------|----------------------------|
| LVI-S | | | | |
| <i>Forceful Dimension</i> | 12 | .83 | .16 | .60 |
| Takes Charge | 4 | .64 | .18 | .35 |
| Declares | 4 | .68 | .25 | .33 |
| Pushes | 4 | .76 | .34 | .56 |
| <i>Enabling Dimension</i> | 12 | .83 | .19 | .61 |
| Empowers | 4 | .70 | .20 | .38 |
| Listens | 4 | .79 | .37 | .59 |
| Supports | 4 | .69 | .22 | .40 |
| SLPI | | | | |
| <i>Models the Way</i> | 6 | .85 | .36 | .72 |
| <i>Inspires a Shared Vision</i> | 6 | .87 | .43 | .74 |
| <i>Challenges the Process</i> | 6 | .87 | .40 | .72 |
| <i>Enables Others to Act</i> | 6 | .87 | .39 | .68 |
| <i>Encourage the Heart</i> | 6 | .92 | .58 | .77 |
| REPUTATIONAL EFFECTIVENESS SCALE | | | | |
| <i>Composite Effectiveness (Eff_Comp)</i> | 3 | .85 | .64 | .75 |

Research Hypotheses

The results of analyses to test the following research questions are reported below:

1. Does the LVI-S verify that there are two leadership factors: forceful leadership and enabling leadership?
2. What proportion of the HD reputational effectiveness can be explained by the six LVI-S sub-dimensions? Are any of the six sub-dimensions statistically significant predictors of HD reputational effectiveness as reported by RAs?
3. To what extent do scores on the LVI-S and the SLPI correlate, providing evidence of convergent construct validity of the LVI-S?
4. Do most HDs underestimate or overestimate their F/E dimension scores in comparison to RAs ratings of the HD's F/E dimension scores?
5. Is there a statistically significant mean difference in effectiveness ratings between HDs who overestimate their F/E dimension scores versus HDs who underestimate their F/E dimension scores?
6. Which regression model better predicts reputational effectiveness scores: Model 1, an additive model where interaction between complementary F/E sub-dimensions is not accounted for, or Model 2, a model that incorporates the focal leader's joint standing scores (versatility) on complementary F/E sub-dimensions?

Research Question 1

The first research question explored whether the LVI-S presented a clearly identifiable two-factor structure where the two factors, forceful and enabling leadership,

are inversely related. The underlying design of the LVI-S was modeled on the executive version of the LVI. Therefore the number of items, thematic content of items, and the factorial analysis of the LVI-S were intended to approximate the psychometric properties and structure of the executive LVI. This research question evaluated whether or not the LVI-S, using a collegiate population in a housing and residence-life context, could approximate or replicate the psychometric characteristics of the Forceful and Enabling dualities first established in the executive version of the LVI.

A principal components analysis with Varimax rotation was performed on the 12 Forceful and the 12 Enabling items comprising the LVI-S. A summary of the components, item loadings, and communalities (h^2) for each dimension are provided in tables (15-16).

Table 16
Component Matrix (with Varimax rotation): Forceful Dimension

| Forceful Items | Component | | | | h^2 |
|----------------|-----------|-----|-----|-----|-------|
| | 1 | 2 | 3 | 4 | |
| Ftc_1 | | | | .81 | .73 |
| Ftc_2 | | | | .69 | .65 |
| Ftc_3 | | | .82 | | .68 |
| Ftc_4 | | | | .66 | .56 |
| Fd_5 | | | .72 | | .65 |
| Fd_6 | | .74 | .46 | | .78 |
| Fd_7 | | .84 | | | .76 |
| Fd_8 | | .52 | | | .43 |
| Fp_9 | .78 | | | | .68 |
| Fp_10 | .85 | | | | .75 |
| Fp_11 | .59 | .42 | | | .63 |
| Fp_12 | .46 | | .46 | | .54 |

Table 17
Component Matrix (with Varimax rotation): Enabling Dimension

| Enabling Items | Component | | | h^2 |
|----------------|-----------|-----|-----|-------|
| | 1 | 2 | 3 | |
| Ee_13 | | .70 | | .64 |
| Ee_14 | | .77 | | .63 |
| Ee_15 | .61 | .45 | | .60 |
| Ee_16 | | .72 | | .55 |
| EI_17 | .58 | | | .54 |
| EI_18 | .79 | | | .65 |
| EI_19 | .64 | | | .55 |
| EI_20 | | | .76 | .67 |
| Es_21 | | | .81 | .69 |
| Es_22 | .53 | | | .52 |
| Es_23 | .79 | | | .66 |
| Es_24 | .54 | | | .47 |

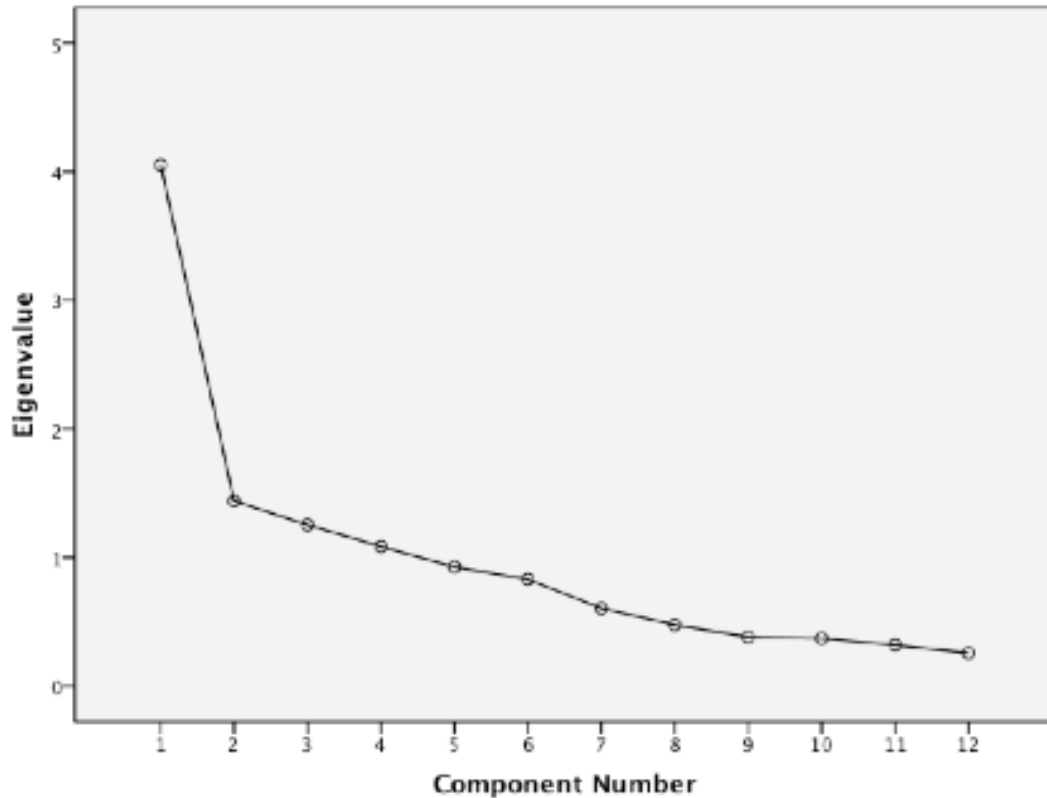
Before interpreting the results of the PCA, a test for sampling adequacy [Kaiser-Meyer-Olkin) (KMO)] and Bartlett’s test of Sphericity were performed. Values on the KMO test of sampling adequacy range from 0 to 1, with values greater than .5 preferred. Both dimensions exceeded the minimum standard (Forceful KMO = .74; Enabling KMO = .77), indicating the sample was satisfactory for a PCA analysis. Bartlett’s test of Sphericity is used to reject the possibility that the correlation matrix is an identity matrix. A significant result ($p < .05$) is required to reject the null hypothesis. The Forceful and Enabling dimensions both produced a significant result (Forceful $p < .01$, Enabling $p < .01$); the null hypothesis was rejected and the PCA results were considered fit for interpretation.

Principal Components Analysis: Forceful Dimension. The PCA analysis on the Forceful dimension (Table 16) extracted four components. The communality (h^2) for each item indicates how well it is represented in the common factor space. The h^2 value for most Forceful items was greater than .5, indicating most items were well represented.

Item Fd_8, “Defends his/her point-of-view—doesn’t back down easily”) produced a communality lower than .5 ($h^2 = .43$). Despite this lower value, Fd_8 was retained because it loaded effectively on component 2 (along with two other items intended to measure the same construct).

A cursory review of Tables 16 and 17 reveal that several variables loaded on their primary component but exhibited secondary loadings elsewhere. Cross-loaded variables, also known as complex variables, often warrant elimination but caution should be heeded in the case of the Leadership Versatility Index. The LVI-S generates predictive validity from the measurement scales as well as through an item-level relationship between the scales. The versatility score (computed through the joint-standing variable) is calculated based on the joint relationship between paired items, their parent sub-dimensions, and the overarching dimensions composing either side of the Forceful and Enabling duality. Eliminating a variable may increase the explained variance and enhance scale alphas. On the other hand, reducing variables can compromise the instrument’s predictive ability because eliminating an item sacrifices not only that item, but also the item’s complement in the opposing dimension. The process for assessing complex variables will be addressed after the results of the PCA have been presented.

Figure 4
Forceful Dimension Scree Plot

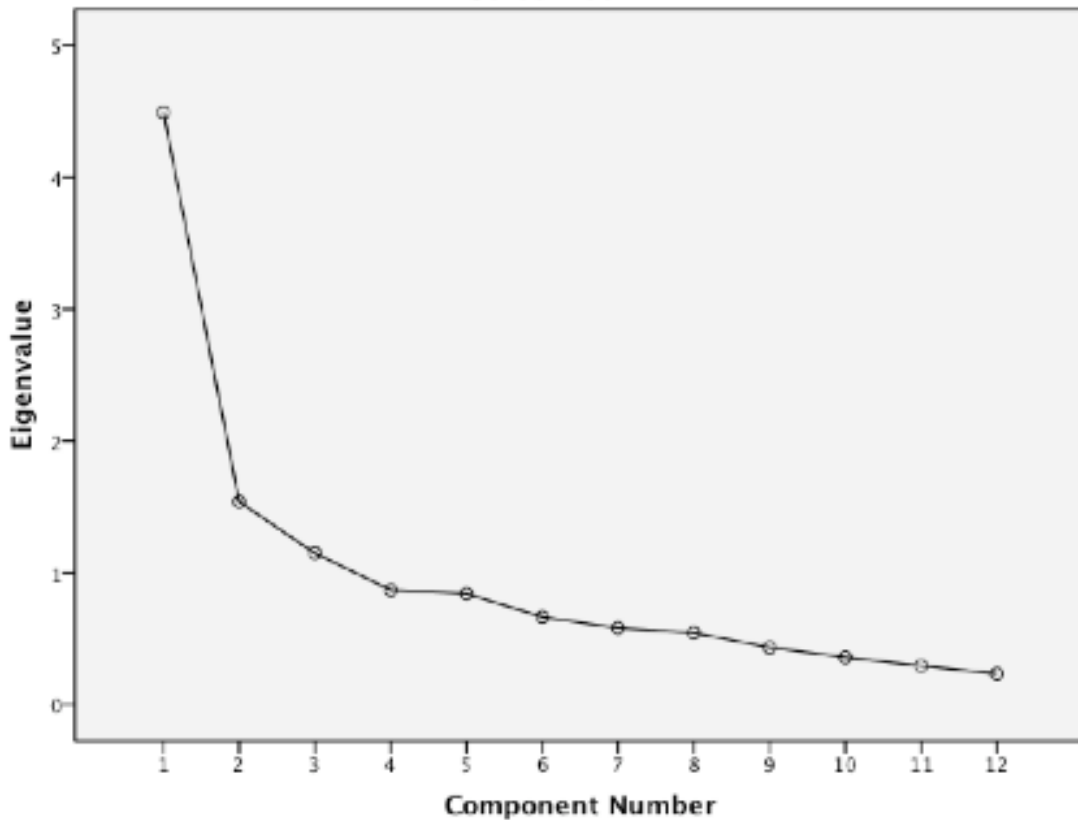


The PCA of Forceful leadership extracted four components that explained approximately 65% of the variance. Component 1 (eigenvalue = 4.05) accounted for 33.7% of this variance. The four items loading on component 1 were intended to measure the *Pushes* construct (Fp_9, Fp_10, Fp_11, and Fp_12), providing empirical support for the *Pushes* sub-dimension. Component 2 (eigenvalue 1.439) explained 12% of the variance and consisted of three items intended to measure the *Declares* construct (Fd_6, Fd_7, and Fd_8). Component 2 provided empirical support for the *Declares* sub-dimension. Component 3 (eigenvalue 1.09) presented the greatest interpretive challenge in the Forceful dimension. It explained 10.5% of the variance but the items loading on

component 3 were intended to measure three different constructs. A common thread among these items was assertive, overt action. Component 4 (eigenvalue 1.09) explained 9% of the variance in Forceful leadership. All three items loaded on component 4 were intended to measure the *Takes Charge* sub-dimension (Ftc_1, Ftc_2, and Ftc_4), therefore component 4 provided empirical support for the *Takes Charge* sub-dimension. All other components had eigenvalues less than 1.0, indicating that the 12 items comprising the sub-dimensions of Forceful leadership were measuring approximately four latent factors. Three of the components aligned with the intended sub-dimension constructs (components 1, 2, and 4), although component 3 appeared to consist of highly assertive elements from each of the three sub-dimensions.

Principal Components Analysis: Enabling Dimension The PCA for the Enabling dimension revealed three components and explained approximately 59% of the variance. Seven items loaded on component 1 (eigenvalue = 4.36) and accounted for 36.3% of the variance in Enabling leadership. Since component 1 had seven items, it was apparent that more than one sub-dimension was significantly represented. Three out of four items intended for measuring the *Listens* sub-dimension were loaded on component 1; three of the four remaining items 1 were from the *Supports* sub-dimension. Overall, it appeared the act of listening was considered analogous with the more general concept of supporting others.

Figure 5
Enabling Dimension Scree Plot



Component 1 provides evidence to support the Listening sub-dimension, or the *Supports* sub-dimension, or a combination of the two. It does not provide evidence that both can be represented as distinct sub-dimensions. Component 2 (eigenvalue = 1.56) accounted for 13% of the variance and was clearly aligned with all four items intended for measuring *Empowers Others*. Component 2 provided evidence in support of the *Empowers Others* sub-dimension. Component 3 consisted of one item each from *Supports* and *Listens* and one cross-loaded item from *Empowers Others* (Ee_13). Validating others appears to be a common thread among Component 3 items: “empowers others,” “makes it safe to challenge his/her thinking,” and “shows appreciation.” As with

Component 3 in the Forceful PCA, this component seems to contain a selection of items occupying space between Supportive behavior (inclusive of Listening) and *Empowers Others*. This result does not contradict, or necessarily confound the pre-determined structure of the Enabling dimension, but it does suggest a need for further refinement of items and, possibly, the need to design a more unique sub-dimension that reflects the personal validation aspects of supportive behaviors.

Inverse relationship of the Forceful and Enabling dimensions. A key aspect of the executive version of the LVI is the inverse relationship between the dimensions and sub-dimensions. These correlations are an important aspect of the model, for if dimensions and sub-dimensions are not negatively correlated then they are not complementary, opposing dimensions or sub-dimensions. A correlation analyses was performed to examine whether the LVI-S Forceful and Enabling dimensions were inversely related to one another. The overall correlation between the Forceful and Enabling dimensions was $-.47 (p < .01)$, indicating that their scales do appear to function in an inverse fashion.

Research question 1 attempted to determine whether the dimensions and sub-dimensions of Forceful and Enabling leadership from the executive LVI could be approximated in a collegiate population. Support was found for five out of six sub-dimensions. One of the sub-dimensions could be more parsimoniously explained by combining the existing *Supports* and *Listens* sub-dimensions. Consideration might be given to re-conceptualizing the third *Supports* sub-dimension to better represent validating or encouraging leadership characteristics.

Research Question 2

The second question explored whether any of the six LVI-S sub-dimensions were statistically significant predictors of HD reputational effectiveness. Two hypotheses were offered. First, RA-reported (observer) sub-dimension scores were expected to be statistically significant predictors of leader effectiveness ratings. Second, HD-reported (self) sub-dimension scores were not expected to be statistically significant predictors of leader effectiveness ratings.

As predicted in hypothesis 2a, RA-reported sub-dimension scores significantly predicted leader effectiveness scores, accounting for 56% of the variance in reputational effectiveness ratings ($R = .75$; $R^2 = .56$). As predicted in hypothesis 2b, HD self-reported sub-dimension scores did not significantly predict effectiveness. Tables (18-20) summarize these results.

Table 18
Variance Explained: HD LVI-S Sub-dimension Scores and Composite Effectiveness Score

| Source | Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------------|-------|------------------|------------|---------------------|----------------------------|
| RA Observer | 1 | .75 ^a | .56 | .49 | 1.55 |
| HD Self | 1 | .37 ^b | .14 | .01 | 2.16 |

a. Predictors: (Constant), *Supports, Pushes, Empowers, Declares, Listens, TakesCharge*

b. Predictors: (Constant), *Supports, Pushes, Declares, Empowers, TakesCharge, Listens*

Table 19
Coefficients^a of LVI-S Sub-dimension Scores and Composite Effectiveness Score

| Rating Source | Sub-dimension | <u>Unstandardized Coefficients</u> | | <u>Standardized Coefficients</u> | <i>t</i> | Sig. |
|---------------|---------------|------------------------------------|------------|----------------------------------|----------|------|
| | | <i>B</i> | Std. Error | Beta | | |
| RA Observer | (Constant) | 14.47 | 4.40 | | 3.29** | .00 |
| | RA_TC | 1.92 | .72 | .47 | 2.66* | .01 |
| | RA_Decl | -.30 | .70 | -.07 | -.43 | .67 |
| | RA_Push | -1.87 | .63 | -.43 | -2.98** | .01 |
| | RA_Emp | -1.55 | .71 | -.29 | -2.18* | .035 |
| | RA_Listn | 3.3 | .66 | .78 | 5.01** | .00 |
| | RA_Supp | -.84 | .89 | -.16 | -.93 | .36 |
| HD Self | (Constant) | 19.93 | 3.76 | | 5.30 | 0.00 |
| | RA_TC | -.79 | .50 | -.28 | -1.58 | .12 |
| | RA_Decl | -.01 | .36 | -.01 | -.027 | .98 |
| | RA_Push | .38 | .34 | .20 | 1.11 | .27 |
| | RA_Emp | .07 | .41 | .03 | .17 | .87 |
| | RA_Listn | -.78 | .47 | -.31 | -1.67 | .10 |
| | RA_Supp | .54 | .35 | .25 | 1.54 | .13 |

a. Dependent Variable: Eff_Comp

*p<.05.

**p<.01.

A follow-up analysis was conducted with regard to hypothesis 2b. Because HD sub-dimension scores were unable to significantly predict effectiveness scores, a more direct analysis was conducted. HD (self-reported) sub-dimension scores were used to predict RA (observer) sub-dimensions scores. Only one of the six regressions, *Supports*, was significant. Results from these analyses are provided in Table 20.

Table 20
 Variance Explained: HD (Self-report) Sub-dim Scores and RA (Observer-report)
 Sub-dim Scores

| Sub-dimension | <i>R</i> | <i>R</i> Square | <i>F</i> |
|---------------------|----------|-----------------|----------|
| <i>Takes Charge</i> | .42 | .17 | 2.31 |
| <i>Declares</i> | .28 | .08 | .91 |
| <i>Pushes</i> | .38 | .14 | 1.82 |
| <i>Empowers</i> | .41 | .17 | 2.22 |
| <i>Listens</i> | .16 | .03 | .89 |
| <i>Supports</i> | .67 | .44 | 8.79* |

Dependent variable: RA (observed) sub-dimension score

* $p < .01$.

Research Question 3

The next research question explored the convergent construct validity of the LVI-S with respect to the Leadership Practices Inventory—Student Version (SLPI), an established, student-focused multi-rater leadership instrument. A correlation analysis examined the relationships between the LVI-S and the SLPI (Table 21). This correlation analysis yielded evidence of convergent and discriminant validity for the LVI-S.

Table 21
 Pearson's *r* Correlation Matrix for LVIS Sub-dimensions, SLPI Scales, and
 Reputational Effectiveness Scale

| | LPI Modl | LPI Visn | LPI Chllng | LPI Enabl | LPI Encrg | Eff Comp |
|--------------|--------------|--------------|---------------|--------------|--------------|--------------|
| LVI TC | .25* | .25 | .33* | .04 | .21 | .18** |
| LVI Decl | -.05 | .08 | .04 | -.21* | -.02 | -.13* |
| LVI Push | .04 | .10 | .105 | -.21* | -.02 | -.12 |
| LVI Emp | .01 | -.02 | -.01 | .17* | .01 | .06 |
| LVI Listn | .42* | .34* | .33* | .54* | .37* | .48** |
| LVI Supp | .32* | .28* | .22* | .43* | .38* | .33** |
| Eff Comp | .70** | .63** | .65** | .70* | .59* | - |

**p* < .01

n = 261

LVI TC = LVI *Takes Charge* Sub-dimension

LVI Decl = LVI *Declares* Sub-dimension

LVI Push = LVI *Pushes* Sub-dimension

LVI Emp = LVI *Empowers* Sub-dimension

LVI Listn = LVI *Listens* Sub-dimension

LVI Supp = LVI *Supports* Sub-dimension

LPI Modl = Models the Way Scale

LPI Visn = Inspires a Shared Vision Scale

LPI Chllng = Challenge the Process Scale

LPI Enabl = Enables Others to Act Scale

LPI Encrg = Encourages the Heart Scale

Eff Comp = Reputational Effectiveness Scale

In hypothesis 3a, the Enabling sub-dimensions of the LVI-S (*Empowers*, *Listens*, and *Supports*) were expected to correlate significantly with the SLPI scales Enabling Others to Act, Inspiring a Shared Vision, and Encouraging the Heart. Hypotheses 3a was supported, providing evidence of convergent validity. *Listens* and *Supports* were positively correlated with all three hypothesized scales. Additionally, significant positive correlations were found for Models the Way and Challenge the Process. The third LVI-S

sub-dimension, *Empowers*, correlated positively and significantly with Enables Others to Act but was not significantly correlated with the other SLPI scales.

In hypothesis 3b, the forceful sub-dimensions of the LVI-S (*Takes Charge*, *Declares*, and *Pushes*) were expected to correlate significantly with the Challenge the Process scale from Kouzes and Posner's SLPI. *Takes Charge* demonstrated a significant, positive correlation with Challenge the Process ($r = .33, p < .01$) providing convergent evidence. Although not hypothesized, *Takes Charge* also exhibited significant positive correlations with Models the Way ($p = .25, p < .01$). *Declares* and *Pushes* were not significantly correlated with Challenges the Process. Negative correlations between Enables Others to Act and the LVI-S sub-dimensions *Declares* ($r = -.21, p < .01$) and *Pushes* ($r = -.21, p < .01$) were interpreted as evidence of discriminant validity for the forceful dimension. Further explanation of this interpretation is provided in Chapter 5.

Research Question 4

The next research question explored whether HDs self-rated forceful and enabling dimension scores were underestimated or overestimated in comparison to RA observer ratings. It was hypothesized the majority of HDs would overestimate their scores in comparison to their observers.

Before assigning rater agreement categories, a series of one-way ANOVA analyses compared HD and RA means for each LVI-S sub-dimension. These analyses helped confirm whether HD means differed significantly from RA means. Three mean comparisons produced significant *F* scores, indicating mean differences between rater groups. When means differ significantly, raw mean scores should not be used to

determine rater agreement categories (Atwater & Yammarino, 1992). Table 22 provides a summary of the LVI-S sub-dimension means comparison analysis. Since HD and RA mean scores differed significantly in three of the six sub-dimensions (*Takes Charge*, *Pushes*, and *Listens*), rater agreement categories were not determined via raw mean scores.

Table 22
Means Comparison of HD and RA LVI-S Sub-dimension Scores

| LVI-S Sub-dimension | F |
|---------------------|---------|
| <i>Takes Charge</i> | 15.95** |
| <i>Declares</i> | .165 |
| <i>Pushes</i> | 5.35* |
| <i>Empowers</i> | 2.12 |
| <i>Listens</i> | 13.25** |
| <i>Supports</i> | 2.60 |

*p<.05.

**p<.01.

Atwater and Yammarino (1992) recommended using difference scores to categorize HDs into agreement categories. Atwater and Yammarino (1992) calculated difference scores for all leader-subordinate comparisons, computed the mean difference score within the distribution, and then categorized rater agreement groups based on this distribution of difference scores. Accordingly, the following process was used to address RQ4. HDs with difference scores within one-half a standard deviation of the mean difference for that comparison were considered “self-aware” because their ratings approximated the scores provided by their raters. Difference scores one-half standard deviation or more above the mean difference for that comparison were considered “over-estimators.” A difference score one-half standard deviation or more lower than the mean

difference for that comparison were considered “under-estimators” (Atwater & Yammarino, 1992). For a more complete accounting of this process, see Atwater and Yammarino (1992; pp. 151-152). Table 23 summarizes the data used for rater assignment categories.

Table 23
Dimension-level Mean Difference Scores and Boundary Parameters for Rater Agreement Categories

| | Mean DiffScore | SD | Lower Bound (Under-estimator) | Upper-Bound (Over-estimator) |
|--------------------|----------------|-----|-------------------------------|------------------------------|
| Forceful Dimension | -0.29 | .45 | < -.52 | > -.06 |
| Enabling Dimension | -.36 | .46 | < -.59 | > -.14 |

Agreement categories were calculated separately for each dimension because some HDs were categorized differently in one dimension than in its complement. Consolidating the analysis across dimensions would confound the results. Table 24 reports the crosstab results for rater assignment categories.

Table 24
Crosstab Analysis: Rater Assignment Categories and Chi-Square Test
Enabling Dimension

| | Under-estimator | Self-aware | Over-estimator | Total |
|---------------------------|-----------------|------------|----------------|-----------|
| Forceful Dimension | | | | |
| Under-estimator | 6 | 3 | 3 | 12 |
| Self-aware | 4 | 13 | 6 | 23 |
| Over-estimator | 3 | 7 | 4 | 14 |
| Total | 13 | 23 | 13 | 49 |

| | Ch-Square Test | | |
|--------------------|----------------|----|-----------------------|
| | Value | df | Asymp. Sig. (2-sided) |
| Pearson Chi-Square | 4.845a | 4 | .30 |
| N of Valid Cases | 49 | | |

a. 4 cells (44.4%) have expected count less than 5. The minimum expected count is 3.18.

Slightly more than half the sample (26 of 49 HDs) categorized themselves differently on one dimension than on the other. Six (12.2%) of these disparities were of the extreme variety (overestimating in one category vs. underestimating in the other); the remaining 20 cases (40.8%) consisted of agreement in one dimension but overestimation or underestimation in the other. A Chi Square test was performed to determine if agreement groups were distributed differently according to Forceful and Enabling dimensions. The test failed to indicate a significant difference, $\chi^2(4) = 4.845, p = .30$ (an alpha level of .05 was utilized).

Hypothesis 4 proposed most HDs would over-estimate their leadership ratings in comparison to those provided by their RAs. This hypothesis was not supported in either the Forceful or Enabling Dimension. Thirteen HDs (26.5%) over-estimated their scores in the Enabling Dimension; of the remainder, 23 (46.9%) were categorized as self-aware and 13 (26.5%) were categorized under-estimators. In the Forceful Dimension, a similar pattern emerged. Fourteen (28.6%) HDs over-estimated their scores, 23 (46.9%) were self-aware and the remaining 12 (24.5%) were categorized as under-estimators.

Research Question 5

Research question 5 explored whether there would be a statistically significant mean difference in effectiveness scores between HDs who overestimated their scores versus HDs who under-estimated their scores. It was further hypothesized that over-estimators would exhibit lower effectiveness scores than HDs who under-estimated their scores. As in research question 4, analyses were conducted separately for the Forceful and Enabling dimensions.

Forceful dimension analysis. The initial analysis consisted of a one-way ANOVA that compared composite effectiveness scores across rater agreement categories for each dimension. Eight analyses were conducted, one for each dimension and six for the sub-dimensions therein. Descriptive statistics and results from the Forceful dimension *F* tests are contained in Table 25.

Table 25
Forceful Dimension Mean Comparisons: Agreement Categories by Composite Effectiveness (Eff_Comp) Score

| Sub-dimension | Agreement Category | N | Mean Eff_Comp | Standard Deviation | <i>F</i> |
|---------------------|--------------------|----|---------------|--------------------|----------|
| Forceful Dimension | Under | 12 | 16.11 | 1.59 | .49 |
| | Aware | 23 | 16.88 | 2.01 | |
| | Over | 14 | 16.71 | 2.67 | |
| <i>Takes Charge</i> | Under | 12 | 15.95 | 2.09 | 5.90* |
| | Aware | 25 | 17.57 | 1.54 | |
| | Over | 12 | 15.4 | 2.58 | |
| <i>Declares</i> | Under | 9 | 16.41 | 2.1 | .12 |
| | Agree | 33 | 16.75 | 2.12 | |
| | Over | 7 | 16.42 | 2.65 | |
| <i>Pushes</i> | Under | 12 | 16.17 | 2.7 | .37 |
| | Aware | 22 | 16.75 | 1.98 | |
| | Over | 15 | 16.85 | 2.02 | |

*p<.01.

Hypothesis 5a proposed that effectiveness scores for over-raters would be significantly different than effectiveness scores for under-raters. The test on Forceful dimension agreement categories was not significant, indicating there were no significant differences across mean scores of effectiveness (Table 25). At the sub-dimension level, however, *Takes Charge* yielded a significant result not reflected in the dimension-level

test ($F = 5.90, p < .05$). A Games-Howell post-hoc analysis determined the direction of mean differences. Levene's statistic was significant ($p < .05$), indicating the variances from the populations were not homogenous. Welch's test for independent samples was applied to provide a more robust standard for comparison and the result remained significant ($p = .01$). The post-hoc analysis indicated the difference was due to the self-aware group's higher mean effectiveness score. Since the over-estimator and under-estimator mean scores were not significantly different, so the *Takes Charge* analysis did not support hypothesis 5a. It should be noted that the internal consistency coefficient for *Takes Charge* is low ($\alpha = .64$), therefore the standard error of measurement is greater than preferred. Consequently this result should be interpreted with caution.

Enabling dimension analysis. Analyses on the Enabling dimension yielded support for hypotheses 5a and 5b. A test at the dimension level was significant ($F = 6.88, p < .01$); this result indicated that effectiveness means were significantly different across agreement categories. The ANOVA mean comparisons for the Enabling dimension and sub-dimension are summarized in Table 26.

Table 26
 Enables Dimension Mean Comparisons: Agreement Categories by Composite Effectiveness (Eff_Comp) Score

| Sub-dimension | Agreement Category | N | Mean Eff_Comp | Standard Deviation | F |
|------------------------|--------------------|----|---------------|--------------------|-------|
| Enabling Dimension | Under | 13 | 17.46 | 1.56 | 6.88* |
| | Aware | 23 | 17.17 | 1.40 | |
| | Over | 13 | 15.03 | 2.8 | |
| <i>Empowers Others</i> | Under | 14 | 16.96 | 1.39 | .38 |
| | Aware | 15 | 16.26 | 2.97 | |
| | Over | 20 | 16.69 | 1.94 | |
| <i>Listens</i> | Under | 14 | 16.71 | 1.70 | 4.85* |
| | Aware | 22 | 17.42 | 1.66 | |
| | Over | 13 | 15.24 | 2.16 | |
| <i>Supports</i> | Under | 14 | 16.9 | 1.63 | 5.60* |
| | Aware | 19 | 17.54 | 1.99 | |
| | Over | 16 | 15.33 | 2.22 | |

* $p < .01$.

A Games-Howell post-hoc test was used to determine the direction of mean differences within the Enabling dimension. Levene's statistic was significant so equal variances were not assumed. Welch's test was applied to provide a more robust standard for comparison and results remained significant ($p = .03$). The post-hoc analysis indicated the over-rater mean score ($m = 15.03$) was significantly different from both under-rater ($m = 17.46$) and in agreement ($m = 17.17$) scores. This finding supported hypothesis 5a. Since the over-rater effectiveness mean score proved significantly lower than the under-rater score, this finding also provided support for hypothesis 5b.

The *Listens* sub-dimension produced a significant ANOVA result ($F = 4.85, p = .01$). As in previous analyses, this sample also yielded a significant Levene's statistic ($p =$

.03) so equal variances were not assumed. Welch's test was applied and the result remained significant ($p < .05$). Post-hoc analyses were conducted to determine which means were significantly different. Significant differences were attributed to the self-aware category ($m = 17.42$). The over-estimator mean ($m = 15.24$) and the under-estimator mean ($m = 16.71$) were not significantly different from one another, so these results did not provide support for hypothesis 5a or 5b.

The *Supports* sub-dimension yielded a significant ANOVA test ($F = 5.60, p = .01$). Levene's test was not significant in this case, so Tukey's HSD post-hoc analysis was applied to determine the direction of mean differences. Means for self-aware ($m = 17.54$) and under-rater ($m = 16.9$) categories were not significantly different, nor were means for the under-rater ($m = 16.9$) and over-rater ($m = 15.33$) categories. Significant differences were found between the over-rater and self-aware means ($p = .005$), however these findings did not support hypotheses 5a or 5b. As noted in the findings from the *Takes Charge* analysis, the *Supports* sub-dimension findings should be interpreted with some caution because the internal consistency coefficient ($\alpha = .69$) is slightly lower than would be preferred interpretation of analyses at the sub-dimension level.

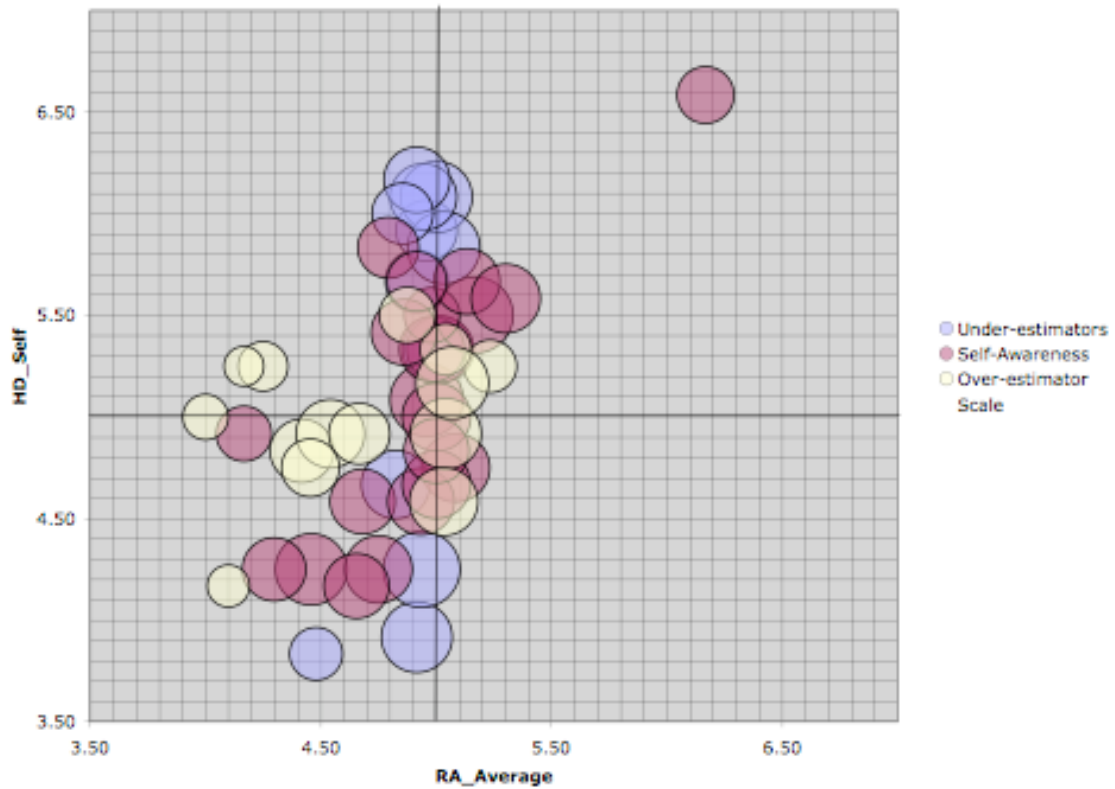
Direction of rating discrepancies. To better understand the nature of these mean differences, the relationship between effectiveness and agreement categories was investigated further. The difference score calculations used to assign agreement categories do not provide insight into whether or not rating discrepancies were due to over-use or under-use of behaviors. For example, one could not determine the proportion of under-estimators who rated themselves less optimally because they over-used enabling

behaviors versus those who under-estimated themselves because they under-used them. The following charts provide insight into the direction of rating discrepancies. Figures (2-5) depict relationships between HD self-ratings, RA observer ratings (averaged), and effectiveness scores within each agreement category. RA ratings were plotted on the X-axis and HD self-scores were plotted along the Y-axis. Marker size denoted the average effectiveness score provided by RAs (Eff_Comp); larger markers indicated higher effectiveness scores.

Two important points should be noted before presenting these charts. First, a numerically higher rating does not necessarily indicate a more optimal rating. The optimal value on the 9-point Too Little, Too Much (TMTL) scale is in the middle (5). To make this easier to observe, each axis was marked with a reference line that extends from 5, indicating “the right amount” on the TLTM scale. Values lower than 5 indicate under-use of enabling behaviors; values higher than 5 indicate over-use. Thus, “over-estimators” provided self-scores that were closer to optimal than those assigned by their observers; the reverse was true for “under-estimators.”

Second, the data plotted in these charts is not identical to the data used to define agreement categories. Agreement categories were calculated from difference scores, but these charts plot self- and observer-rated mean scores. Close inspection reveals cases where self-rated scores and observer scores appear nearly identical, yet the HD is categorized as an under- or over-estimator. In these cases, the sub-dimension means for self- and observer-scores differed substantially, but the differences were cancelled out in the dimension-level mean scores.

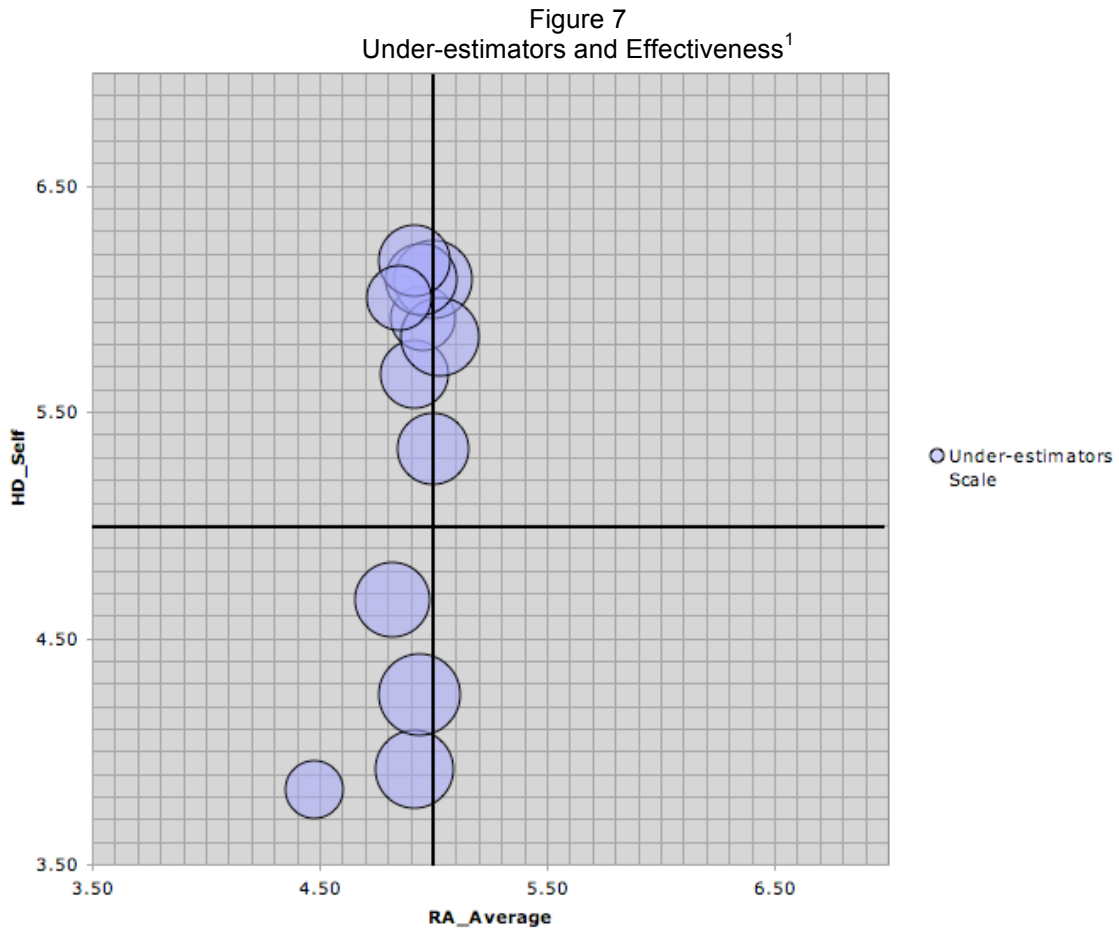
Figure 6
Rater Agreement Categories and Effectiveness¹



¹ Size of circle represents effectiveness rating: larger = higher rating

Figure 6 shows all agreement categories and their relative position to one another. A majority of RA_average ratings clustered on or near the optimal value. Self-aware ratings hinted at a diagonal trend, under-estimator ratings presented a distinct vertical orientation and the over-estimator ratings displayed a more horizontal and leftward orientation. Each agreement category has unique features and characteristics that are difficult to observe when all categories are combined in one chart. To ease interpretation, each agreement category has been portrayed in individual charts (Figures 7-9).

To explore the effects of rating incongruity, the under-estimator (Figure 7) and over-estimator (Figure 8) charts will be presented first. A final chart (Figure 9) will address self-awareness and effectiveness ratings.

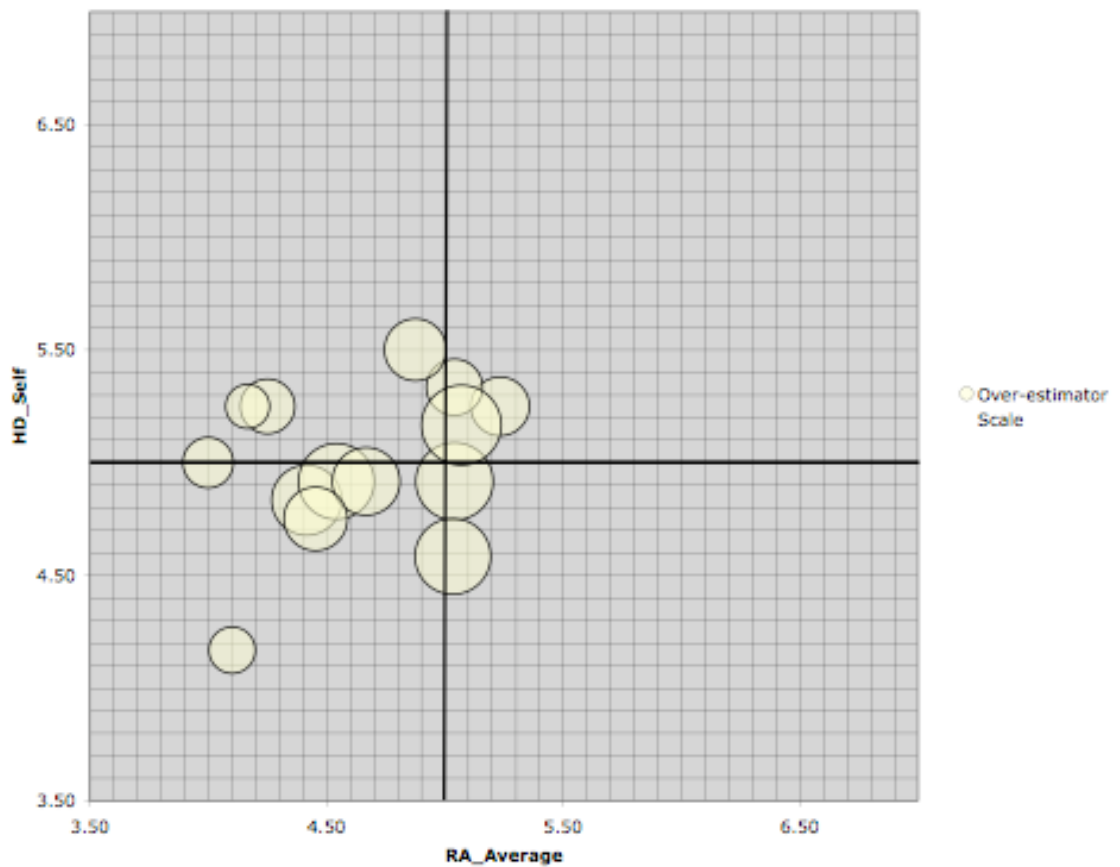


¹ Size of circle represents effectiveness rating: larger = higher rating

Under-raters scored themselves less optimally than did their observers. Since the majority of RA ratings clustered at or near the middle of the scale, a clear division can be seen between the two types of under-rating. Eight HDs rated themselves less optimal because they believed they over-used enabling behaviors (top portion of the chart). Four

HDs rated themselves less optimal because they believed they were not enabling enough in their leadership. For the most part, neither form of under-rating appeared to impact effectiveness in a detrimental way. Most under-estimators received scores that comparable to those in the self-aware category.

Figure 8
Over-estimators and Effectiveness¹



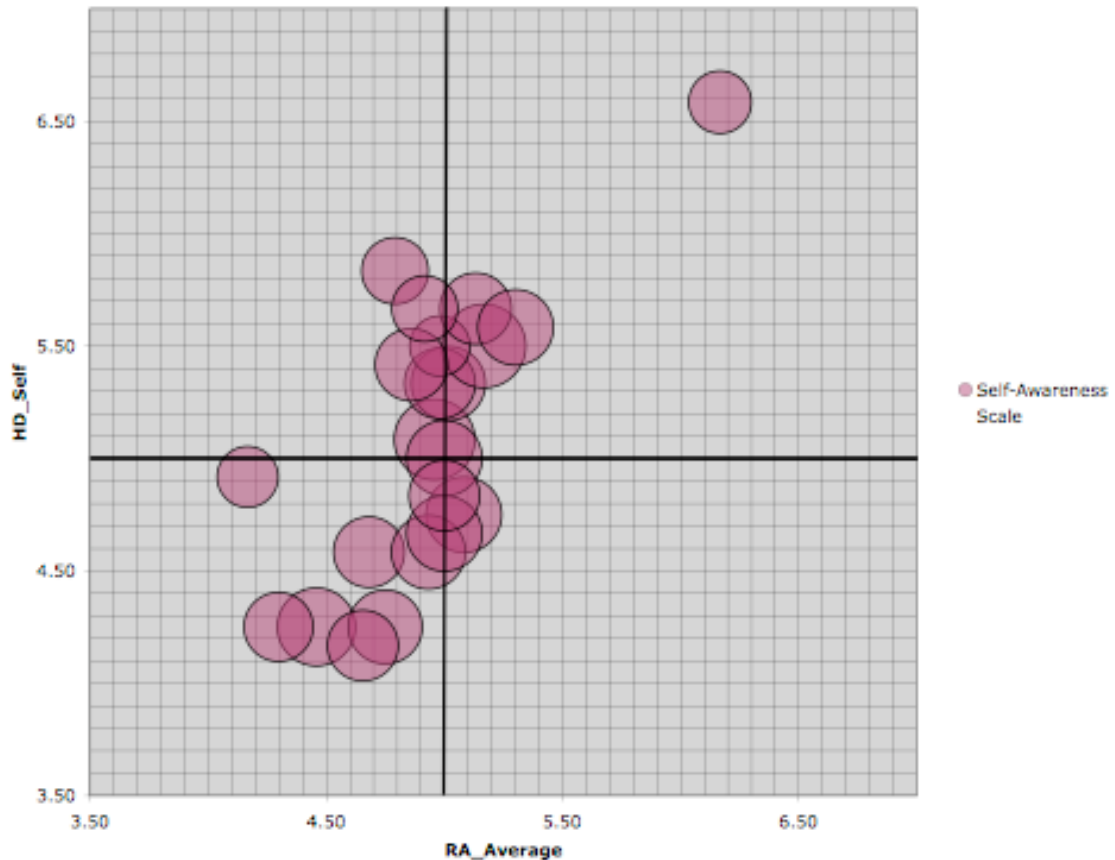
¹ Size of circle represents effectiveness rating: larger = higher rating

Rating incongruity is not always benign. Compared to other agreement categories, over-estimator self-ratings exhibited less range. In the other groups (comprised of 35 self-raters), only 4 self-ratings were within two tenths of the ideal (approximately 11.4%).

Over-estimator self-scores fell within a half-point of the ideal value (between 4.5 and 5.5) and approximately 50% were within two tenths of the ideal rating (8 out of 14). Over-estimators not only perceived their leadership in a more favorable light than their RAs; as a group, over-estimators rated themselves more favorably than other HDs.

Over-estimation influenced effectiveness scores, especially for HDs plotted on the left (too little) side of the RA_average reference line. Most of these HDs believed they used the right amount (or perhaps too much) enabling behavior, but RAs disagreed. Of all the agreement categories, over-estimators received the lowest effectiveness scores.

Figure 9
Self-Awareness and Effectiveness¹



¹ Size of circle represents effectiveness rating: larger = higher rating

The self-aware agreement category displayed the widest range of values. The majority of ratings grouped close to the RA_average reference line, but notable exceptions were present. Several cases of self-aware over-doing and under-doing were represented. Eight cases of under-doing enabling behaviors were observed and four cases of over-doing were noted. As can be seen in Figure 9, one over-doing case deviated from the reference lines by a wider margin than any other plotted point. The effectiveness score for this individual, nor the others, did not appear to suffer greatly. Several of them were slightly lower than those closer to the ideal, but self-aware deviations were not affected nearly as severely as those in the over-rater category.

Enabling dimension agreement groups and correlates of effectiveness. The next series of analyses explored the correlates of effectiveness within the Enabling dimension. Table 27 reports Enabling dimension means and from two rating sources (observer-rated and self-rated). It presents RA-reported effectiveness scores (Eff_Comp) and effectiveness correlations for RA scores as well as effectiveness correlations for HD scores. The data are parsed four ways. First, a combined category explores effectiveness correlates independent of rater agreement group. The other three categories present correlates as they relate to agreement category. Table 28 presents the same data format for the Enabling sub-dimensions.

Table 27
Enables Dimension Mean Comparisons: Agreement Categories and Correlates of Effectiveness by Rater Source

| | Agreement Category (<i>n</i>) | <u>Rater Source and Effectiveness Means</u> | | | <u>Rater Source and Effectiveness Correlations</u> | |
|--------------------------------------|---------------------------------|---|------|----------|--|-----------|
| | | RA | HD | Eff_Comp | RA (observer) | HD (self) |
| Enabling Dimension $\alpha = .83$ | <i>Combined (49)</i> | 4.83 | 5.15 | 16.64 | .46** | .035 |
| | Under (13) | 4.89 | 5.46 | 17.46 | .66* | -.321 |
| | Aware (22) | 4.92 | 5.08 | 17.17 | .02 | -.069 |
| | Over (14) | 4.63 | 4.99 | 15.03 | .61* | .001 |

* $p < .05$.

** $p < .01$.

At the combined level, RA-provided dimension scores and effectiveness scores are significantly and positively correlated ($r = .46, p < .01$). Within agreement categories, however, significant correlations are not shared across each category. Under-rater and over-rater categories indicate significant correlations but the self-aware group showed virtually no discernable relationship. Self-rated dimension scores exhibited no significant relationships.

Table 28
Enables Sub-dimension Mean Comparisons: Agreement Categories and Correlates
of Effectiveness by Rater Source

| | Agreement Category (n) | <u>Rater Source and Effectiveness Means</u> | | | <u>Rater Source and Effectiveness Correlations</u> | |
|---|---------------------------|---|------|----------|--|--------------|
| | | RA | HD | Eff_Comp | RA (observer) | HD (self) |
| <i>Empowers Others</i> Sub-dim $\alpha = .70$ | Combined (49) | 5.00 | 5.2 | 16.64 | .08 | .03 |
| | Under (14) | 4.93 | 5.73 | 16.96 | -.05 | .230 |
| | Aware (15) | 4.87 | 4.76 | 16.26 | .55* | -.04 |
| | Over (20) | 5.14 | 5.16 | 16.69 | -.21 | -.59** |
| <i>Listens</i> Sub-dim $\alpha = .79$ | Combined (49) | 4.65 | 5.16 | 16.64 | .59** | -.17 |
| | Under (14) | 4.85 | 5.35 | 16.71 | -.28 | -.44 |
| | Aware (22) | 4.80 | 5.11 | 17.42 | .65** | .25 |
| | Over (13) | 4.09 | 5.07 | 15.24 | .80** | -.35 |
| <i>Supports</i> Sub-dim $\alpha = .69$ | Combined (49) | 4.85 | 5.10 | 16.64 | .42** | .12 |
| | Under (14) | 4.87 | 5.34 | 16.9 | .54* | .16 |
| | Aware (19) | 4.96 | 5.14 | 17.54 | .15 | .08 |
| | Over (16) | 4.71 | 4.84 | 15.33 | .51* | .00 |

*p<.05.

**p<.01.

A sub-dimension analysis (Table 28) also demonstrated differences between the combined scores and those organized by agreement category. In the *Empowers Others* sub-dimension, the combined level correlation between observer ratings and effectiveness was nearly non-existent ($r = .08$), but the self-aware agreement group produced a significant correlation ($r = .55$, $p < .05$) that would have gone unnoticed if the data were not parsed. The *Listens* sub-dimension produced strong significant correlations at the combined level ($r = .59$, $p < .01$), but these correlations were even stronger within the self-aware ($r = .65$, $p < .01$) and the over-estimator ($r = .80$, $p < .01$) categories.

Research Question 6

In research questions 1-5, LVI-S dimension and sub-dimension scales measured leadership behaviors independently of one another. In other words, the LVI-S has demonstrated the ability to measure leadership behaviors, but it has not yet measured leadership versatility. Kaplan and Kaiser (2003) conceptualized versatility as the ability to pivot from one style or approach to another depending on the needs of a situation. Kaplan and Kaiser (2003; 2006) presented leadership versatility as a dynamic relationship between seemingly opposite virtues. For the LVI-S to measure versatility, the instrument needed to connect and measure “opposing virtues.”

The LVI-S items, sub-dimensions, and dimensions on each side of the Forceful / Enabling duality mirror one another in a complementary, but opposing fashion at the item level, the sub-dimension level, and the dimensional level (a visual depiction of these relationships is provided in Appendix Z). The instrument measured versatility by considering a leader’s joint standing on complementary opposite behaviors. The “joint-standing” score calculated how effectively a leader chose between Forceful and Enabling approaches to leadership. Joint-standing scores were computed for all complementary item pairs, sub-dimension pairs, and dimension pairs within the instrument. They provided a quantitative measure of versatility, ranging from 0 to 5.66. A more detailed explanation of these computational processes has been provided in Chapter 3 and in Appendix X. Low joint standing scores indicated leaders were perceived to use “the right amount” of Forceful and Enabling behaviors, signifying high levels of versatility.

Conversely, high joint standing scores indicated leaders scored farther away from “the right amount,” signifying lopsided leadership tendencies.

Research question 6 compared the predictive ability of two regression models. An “additive” model predicted effectiveness using the six LVI-S sub-dimensions in a sequential regression. This model did not account for any complementary interaction between the LVI-S scales. A second “versatility” model predicted effectiveness through joint-standing variables. The versatility model accounted for interactions between complementary sub-dimensions. Hypothesis 6 proposed the versatility model would out predict the additive model.

Before presenting the results, some significant differences should be noted between the effectiveness variable used in research question 6 (Eff_ind) versus the one used in research question 2 (Eff_comp). Mono-source bias was an acknowledged limitation of the results from question 2. Podsakoff and Organ (1986) cautioned against research designs that incorporated behavior ratings and effectiveness ratings from the same source. Mono-source designs can lead to artificially high relationship coefficients. A within source, split-sample technique was used to control for mono-source bias in research question 6 (Kaiser, 2006).

In the split-sample technique, behavior ratings were provided by one source and the effectiveness ratings were provided by another source. To accomplish this objective, all observers within a rater group were randomly split into two separate sub-groups (a) and (b). Average behavior ratings and effectiveness ratings were calculated for each sub-group. The ensuing regression calculations took the behavior ratings from one sub-group

and predicted the effectiveness ratings from the other sub-group. Thus, the behavior ratings from sub-group (a) were used to predict effectiveness ratings from sub-group (b). Conversely, behavior ratings from sub-group (b) were used to predict effectiveness ratings from sub-group (a). The result of this process produces the split-sample, within source correlation (Table 29).

Second, the sample used for the within source, split-sample technique was the same population that was used for the ICC(k) reliability test. This population had a minimum ratio of 5 raters-per-leader. Splitting the sample into sub-groups necessitated the use of a sample with a higher raters-per-leader ratio.

The initial test, referred to as the additive model, used sub-dimension variables to predict effectiveness. The first regression incorporated all six sub-dimension scales, providing insight into how well the scales predict at the dimension level. Three subsequent regressions were performed, each consisting of two sub-dimension pairings, to test how well the scales predicted at the sub-dimension level. The additive model examined how well Forceful and Enabling sub-dimensions predicted effectiveness without taking complementary interactions into account. The first regression incorporated all six Forceful and Enabling sub-dimensions (R^2 [Forceful and Enabling dimensions] = .31, $F = 2.87$, $sig = .02$). Follow-up regressions with sub-dimension pairings produced the following results: (R^2 [*Takes Charge* and *Empowers Others*] = .13, $F = 3.3$, $sig = .05$); (R^2 [*Declares* and *Listens*] = .13, $F = 3.18$, $sig = .05$); (R^2 [*Pushes* and *Supports*] = .06, $F = 1.28$, $sig = .29$). Regression results and accompanying F -tests for the additive model are summarized in Table 29.

The second regression model explored how effectively versatility predicted effectiveness. These regressions predicted effectiveness using joint-standing variables. Joint standing scores reflect the interaction between complementary item pairs, sub-dimension pairs, and dimension pairs. Data from the versatility regression model are summarized in Table 29 as well. The first analysis explored how well all Forceful and Enabling sub-dimension pairs predicted effectiveness (R^2 [Forceful and Enabling All sub-dims Versatility] = .36, $F = 7.723$, $sig < .01$). A series of follow-up of regressions analyzed how well sub-dimension versatility predicted effectiveness. Sub-dimension pairings were identical to those utilized in the first test: (R^2 [*Takes Charge* and *Empowers Others* Versatility] = .30, $F = 19.26$, $sig < .01$); (R^2 [*Declares* and *Listens* Versatility] = .14, $F = 6.9$, $sig = .01$); (R^2 [*Pushes* and *Supports* joint-standing] = .01, $F = 2.08$, $sig = .48$).

Table 29
Additive and Versatility Model Comparison (Split Sample) Predicting Effectiveness

| Predictors | <i>R</i> | <i>R</i> ² | Adj. <i>R</i> ² | Std. Error | F-Test |
|---|----------|-----------------------|----------------------------|------------|---------|
| Forceful and Enabling Dimensions (All sub-dims Additive) | .55 | .31 | .20 | 1.85 | 2.87* |
| Forceful / Enabling Duality (All Sub-dims Versatility) | .60 | .36 | .31 | 1.72 | 7.72** |
| <i>Takes Charge and Empowers Others</i> (Additive) | .37 | .13 | .09 | 1.97 | 3.30 |
| <i>Takes Charge / Empowers Others</i> (Versatility) | .55 | .30 | .29 | 1.75 | 19.26** |
| <i>Declares / Listens</i> (Additive) | .36 | .13 | .09 | 1.98 | 3.18 |
| <i>Declares / Listens</i> (Versatility) | .37 | .14 | .12 | 1.95 | 6.85* |
| <i>Pushes / Supports</i> (Additive) | .24 | .06 | .01 | 2.06 | 1.28 |
| <i>Pushes / Supports</i> (Versatility) | .11 | .01 | - | 2.08 | .519 |

*p<.05.

**p<.01

Table 30
Variance Explained: SLPI Scales

| Predictors | <i>R</i> | <i>R</i> ² | Adj. <i>R</i> ² | Std. Error | F-Test |
|---------------------------------|----------|-----------------------|----------------------------|------------|---------|
| All 5 Scales (Additive) | .58 | .34 | .25 | 1.79 | 4.05** |
| <i>Models the Way</i> | .51 | .26 | .25 | 1.8 | 15.70** |
| <i>Inspires a Shared Vision</i> | .51 | .26 | .24 | 1.8 | 15.73** |
| <i>Challenges the Process</i> | .55 | .31 | .29 | 1.74 | 19.36** |
| <i>Enables Others to Act</i> | .50 | .25 | .235 | 1.81 | 14.81** |
| <i>Encourages the Heart</i> | .32 | .10 | .083 | 1.98 | 5.10* |

*p<.05.

**p<.01.

Hypothesis 6 proposed the joint-standing model would predict reputational effectiveness more effectively than the additive model. This hypothesis was supported. The Forceful and Enabling additive model, which incorporated all six of the LVI-S scales, accounted for 31 percent of the variance in reputational effectiveness. This regression was significant ($p = .02$). The Forceful and Enabling versatility regression was significant as well ($p < .02$) and explained 36% of the variance in reputational effectiveness.

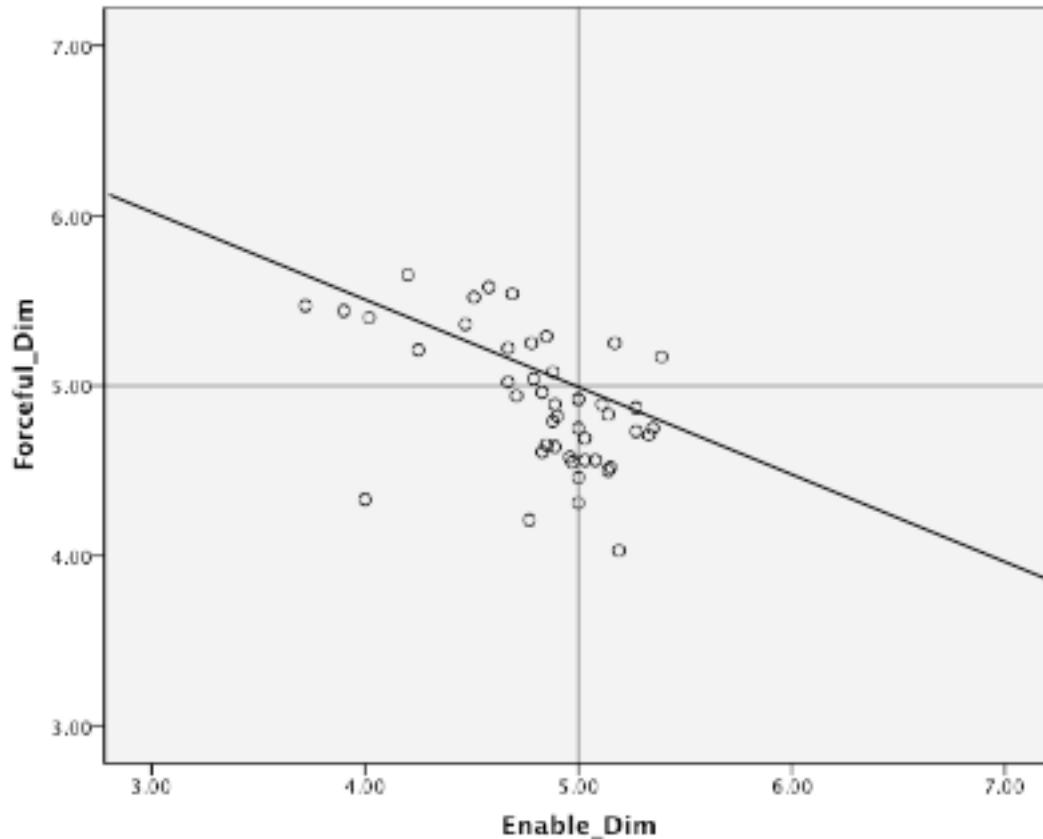
Support for Hypothesis 6 emerged at the sub-dimension level as well. A comparison of the two models (summarized in Tables 29) indicated a difference in predictive power. Only one of three sub-dimension pairings was significant in both the versatility model and the additive model (*Takes Charge / Empowers Others*). The *Takes Charge* and *Empowers Others* versatility regression ($R = .55$, R Square $.30$, $p < .01$), however, explained considerably more variance than the additive model ($R = .37$, R Square $.13$, $p < .05$). The *Declares* and *Listens* pairings were much closer: versatility model ($R = .37$, R Square $= .14$, $p < .05$) and the additive model ($R = .36$, R Square $= .13$, $p > .05$). The *Pushes / Supports* pairings were not significant in either the additive model or versatility model.

On three out of four versatility model regressions, a significant relationship between joint-standing scores and effectiveness was identified. Joint standing scores were calculated based on the relative standing of matched pairs of Forceful and Enabling ratings. The closer ratings were to the response scale's central value ("the right amount"), the lower the joint-standing score. As responses diverged from ideal (either "too little" or

“too much”), joint-standing scores increased. Significant relationships found on three out of four regressions validated the predictive ability of versatility in this population; however, these regressions did not provide insight into implications of different types of rating combinations.

The LVI-S instrument design yields five theoretical categories of variability in joint-standing scores: Under-do Forceful / over-do Enabling (quadrant I); over-do Forceful, over-do Enabling (quadrant II); over-do Forceful / under-do Enabling (quadrant III); under-do Forceful / under-do Enabling (quadrant IV); and Versatile (central position). Only four categories are viable, though, because quadrant II is not realistically possible (Kaplan & Kaiser, 2006). One cannot simultaneously over-do complementary, opposite behaviors. For example, you cannot talk too much and listen too much at the same time.

Figure 10
Distribution of Forceful and Enabling Versatility Plots



n = 66 (33 HDs in split-file format)

Figure 10 indicates a couple of cases rated within the null set. This is most likely due to inexperienced raters and the effect of smaller rater groups in the split sub-sample. Generally speaking, such ratings would average out into a score in one of the other three quadrants (Kaplan & Kaiser, 2006). The scatterplot shows strengths have been over-used in this sample, but not as much as might be expected. The average ratings of overdoing are greater than five, but less than six in most cases. A diagonal regression line shows the characteristic pattern of the polarity effect. When leaders are lopsided, they will bias

themselves in favor of one side of the duality and against the other (Kaplan & Kaiser, 2003a).

CHAPTER V

CONCLUSIONS

Overview

In this chapter, the findings of the study are described. In addition, limitations of the study, implications for college counseling, developing leaders within a campus context, and suggestions for future research are included.

Leadership is not the product of an individual, but rather the result of collaborative group processes. That is, leadership is a social phenomenon. When leader-member relationships are an emphasis of study, these relationships should be measured with instruments capable of assessing them from multiple perspectives, not merely from self-report sourced data (Antonakis, Cianciolo, & Sternberg, 2004). Self reported data are not correlated with objective measures nor with observer reports (Beehr, Ivanitskaya, Hansen, Erofeev & Gudanowski, 2001; Conway & Huffcut, 1997). Observer reports have demonstrated significantly greater validity and reliability than self-reported data (Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988). Therefore leadership research in campus settings should include multi-rater methods that incorporate observer-sourced behavioral feedback.

To better understand Hall Directors' (HDs) leadership behavior and their leader / member relationships with Resident Advisors (RAs), multi-rater instruments should be incorporated into the assessment process (Winston & Fitch, 1993). Kouzes and Posner's

(2003) Leadership Practices Inventory Student Version (SLPI) is a multi-rater instrument validated on a student population. The SLPI has been designed with care and studies have demonstrated it to be a reliable measure of transformational leadership. The SLPI has shortcomings, however, that limit its potential for developing leaders in a campus-housing context.

The SLPI measures relational aspects of leadership such as supportiveness, inclusiveness, and inspiring a common vision—characteristics commonly found in transformational leadership instruments. Although transformational characteristics are critical for residence hall supervision, directive characteristics are also important. The SLPI provides limited measurement of these. Hall Directors must sometimes use their authority to provide direction, establish clear expectations, and hold others accountable to those expectations. But such behaviors should be used with care and intention. Effective residence hall leaders dynamically balance directive and inclusive approaches to leadership.

Winston and colleagues designed the synergistic supervision model to help paraprofessional and professional student affairs supervisors understand the importance of using a “synergistic” leadership style (Winston, Ullum, & Werring, 1984; Winston & Creamer, 1997; 1998; Winston & Hirt, 2003). For example, effective supervisors recognize and address the interpersonal needs of a supervisee, but they do their best to ensure these needs are met without compromising the organizational objectives of the institution (Armino & Winston, 2001; Winston & Creamer, 1997; 1998; Winston & Hirt, 2003). Synergistic supervisors provide structure and direction relative to institutional

responsibilities, but remain open to subordinate perspectives. Supervisors communicate high levels of expectation, but also convey a sense of encouragement and support (Armenio & Winston, 2001). The synergistic supervision model articulates a sophisticated, integrated perspective of leadership (Winston & Creamer, 1997; Winston & Fitch, 1993; Winston, Ullom, & Werring, 1984); however the model does not provide a practical method for measuring synergistic leadership.

To become more synergistic, supervisors must learn to be versatile. They need to develop awareness of the behaviors they rely on too much and those they unwittingly avoid. The SLPI measures a narrow range of constructs that primarily focus on relational, transformational aspects of leadership. It measures these constructs with a frequency scale that is unable to measure overused strengths. Because of these design characteristics, the SLPI is limited in its ability to help HDs become more versatile, synergistic leaders.

The overarching purpose of the current study was to construct and validate a new, multi-rater, student leadership assessment instrument named the Leadership Versatility Index-Student (LVI-S). It was based on an executive version of the Leadership Versatility Index (LVI), an instrument that measures how effectively leaders negotiate the balance between complementary, but opposed leadership behaviors. A student-focused version of the Leadership Versatility Index would be a better match for developing synergistic supervisors.

A literature review explored elements of the versatile leader model and discerned areas of alignment between leadership versatility and synergistic supervision. This review

demonstrated how leadership versatility was relevant to effective leadership in paraprofessional as well as professional roles in student affairs. The methodology section outlined procedures taken to modify the executive-focused LVI (Kaplan & Kaiser, 2006) for use in campus settings. The result of these modifications produced the initial version of the Leadership Versatility Index-Student (LVI-S) that was administered to a sample of Hall Directors (focal leaders) and Resident Assistants (subordinates). The instrument consisted of two separate, but parallel survey forms. The LVI-S “instrument” refers to both the self-rating form and the observer rating form. The “self” version of the survey asked the HD to rate her or himself. The “observer” survey form asked a parallel set of questions where subordinates (RAs) were asked to rate their perception of the HD’s leadership behavior. The data collected by these forms were analyzed according to six research questions:

1. Does the LVI-S verify that there are two leadership factors: forceful leadership and enabling leadership?
2. What proportion of the HD reputational effectiveness can be explained by the six LVI-S sub-dimensions? Are any of the six sub-dimensions statistically significant predictors of HD reputational effectiveness as reported by RAs?
3. To what extent do scores on the LVI-S and the SLPI correlate, providing evidence of convergent construct validity of the LVI-S?
4. Do most HDs underestimate or overestimate their Forceful / Enabling leadership sub-dimension scores in comparison to RAs ratings of the HD’s Forceful / Enabling leadership sub-dimension scores?

5. Is there a statistically significant mean difference in effectiveness ratings between HDs who overestimate their F/E dimension scores versus HDs who underestimate their F/E dimension scores?
6. Which regression model better predicts reputational effectiveness scores: (1) an additive model where interaction between complementary Forceful / Enabling sub-dimensions is not accounted for or (2) a versatility model that incorporates the focal leader's joint standing on complementary Forceful / Enabling sub-dimensions?

The development and validation of the LVI-S provides housing and residence life administrators a tool that measures leadership versatility, a key component of the synergistic supervision model. Such an instrument can provide valuable insights that can inform the development of HD training and supervision programs. Effective programs can help HDs conduct their duties more effectively (Upcraft & Pilato, 1982; Winston & Fitch, 1993) and high quality training was identified as a factor that minimizes RA burnout (Paladino, et al., 2005). Furthermore, an instrument such as the LVI-S can help housing and residence life administrators establish pre and post benchmarking metrics to demonstrate whether training programs are producing their intended effects.

Discussion

Estimates of Reliability

This study provides evidence that the LVI-S is a reliable leadership measure, but further development is warranted. Psychometric testing is an inexact science and some measurement error is expected. Reliability is an estimate of measurement error. The reliability of the LVI-S has been measured through three different statistics: Chronbach's

alpha, Intra-class Correlation Coefficient (*ICC*), and r_{wg} . Chronbach's alpha measured internal consistency, *ICC* measured inter-rater agreement, and r_{wg} measured the degree of interchangeability among raters.

Internal consistency. Chronbach's alpha addressed the internal consistency of raters' responses. The minimum acceptable standard of internal consistency for a psychological measure in the early stages of development is .70 (Nunnally, 1978). This value indicates that 70% of the variance is true score variance, or conversely, that 30% of the variance is due to random measurement error (LeBreton, Burgess, et al., 2003). Higher standards are typically applied to measures in later stages of development, with .80 typically representing the higher threshold. (LeBreton, Burgess, et. al., 2003). The Forceful and Enabling dimensions of the LVI-S produced alpha values that exceeded .80 (Forceful $\alpha = .83$, Enabling $\alpha = .83$). Although dimension level alpha scores exceeded the minimum acceptable standard for a developmental instrument, some sub-dimensions produced coefficients below .70. Lower sub-dimension alphas are understandable—a 4-item scale is more susceptible to variability than a 12-item scale—but strengthening sub-dimension alphas could improve the psychometric performance of the LVI-S. Sub-dimensions were examined to see if elimination of any items would improve the alpha, but in all but one case, the current item structure resulted in the most optimal alpha. The lone exception was item *Ftc_3*, from the *Takes Charge* sub-dimension. With respect to the Chronbach's alpha coefficient, removal of *Ftc_3* would have increased the alpha from .64 to .65. As noted earlier, eliminating an item compromises the balance between the

Takes Charge sub-dimension and the *Empowers Others* sub-dimension. The item was maintained, but it will serve as a primary focus for future research and development.

In comparison to the LVI-S, each of the 5 SLPI scales exhibited stronger internal consistency characteristics. The 6-item scales produced internal measures ranging from .85 to .92. These alphas were consistent with prior research, although slightly higher (Posner, 2005). The relatively large sample size may have contributed to this slight increase.

Inter-rater reliability. The Intraclass Correlation Coefficient (*ICC*) is a measure of inter-rater agreement. Bartko and Carpenter (1976) recommended using intraclass correlation coefficients when ratings are measured on a quantitative scale. *ICC* values range from 0 to 1.0, with .70 serving as the preferred cutoff between a high and low value (LeBreton, Burgess, et. al., 2003). To obtain a high *ICC* value, the scale must exhibit substantial rating consistency (similarity in pattern fluctuation) as well as substantial rating consensus, or similarity in the absolute value of ratings (LeBreton, Burgess, et. al., 2003).

Two versions of *ICC* were reported in this study. The first version *ICC*(1,1) indicated whether individual raters provided information about targets that was reliable and consistent with other raters. Values for *ICC*(1,1) typically range from .00 to .50 with an average of .12 (James, 1982). Conway and Huffcut (1997) reported *ICC*(1,1) values ranging from .2 to .45 in multi-rater studies conducted in a professional context but no multi-rater data on *ICC* scores for college student populations was available. In the current study, *ICC*(1,1) values on the LVI-S ranged from .16 to .37. The *ICC* results for

the LVI-S were acceptable for an instrument under development. The SLPI generated $ICC(1,1)$ values that ranged from .35 to .58, indicating that individual respondents interpreted the SLPI scales more consistently and with greater consensus than the LVI-S scales.

The second version, $ICC(k)$, provided insight into the consistency of mean ratings for a group of k raters. LeBreton, Burgess, et. al., (2003) quoted Bartko's (1976) summarization, "If another random sample of raters rate the same subjects, $ICC(1,k)$ is approximately equal to the correlation between the averaged ratings between the two sets of raters" (p. 89). Essentially $ICC(1,k)$ is the equivalent of the Spearman-Brown prophecy formula applied to $ICC(1,1)$, where k (the number of raters) serves as the correction factor (LeBreton, Burgess, et. al., 2003). James (1982) noted it was possible to take inconsequential values of $ICC(1,1)$ and obtain noteworthy values of $ICC(1,k)$. Adding raters (k) increases the value of the coefficient, therefore $ICC(1,k)$ provides insight into the minimum number of raters needed to obtain a sufficiently high level of reliability on a multi-rater measure.

In the current study, $k = 5$ was determined to provide the greatest value of k with respect to optimizing the number of rater groups available for analysis. In all, 24 rater groups contained 5 or more raters, $n = 160$. The question at hand was whether $k = 5$ can produce a sufficient level of inter-rater reliability for the LVI-S. According to Nunnally (1978), the preferred cutoff value for $ICC(k)$ is .70, though values between .5 and .7 are typically considered acceptable (Kaiser & Kaplan, 2006). ICC scores for LVI-S dimensions were both below the preferred value of .70 but within the range of acceptable

practice (Forceful $k = 5$, .60; Enabling $k = 5$, .61). This result indicated the dimensions provided an acceptable level of inter-rater reliability provided at least 5 observers contribute ratings, but additional raters are recommended to optimize inter-rater reliability.

In comparison to the LVI-S, the five SLPI scales produced *ICC* (1,5) values that ranged from .68 to .77, with four of the SLPI scales yielding coefficients exceeding .70. Given the same number of raters ($k = 5$), the SLPI will provide greater inter-rater reliability. The SLPI scales are more refined, as evidenced by their greater internal consistency and *ICC* coefficients. The SLPI requires fewer raters to obtain comparable levels of inter-rater reliability to the LVI-S, however both instruments provide sufficient levels of inter-rater reliability.

The psychometric properties of the LVI-S would be enhanced if *ICC* scores could meet or exceed .70. Inter-rater agreement and internal consistency are closely related. If a scale measures a latent construct with greater consistency, then it stands to reason multiple raters' scores will exhibit greater consistency and consensus. Thus, as internal consistency (α) improves, so does inter-rater agreement (*ICC*). Refining the items comprising LVI-S sub-dimensions will enhance internal consistency; these improvements should have a positive impact on inter-rater agreement as well.

Measuring Leadership Versatility in a Housing and Residence Life Context

Housing and residence-life administrators strive to help professional and paraprofessional supervisors develop greater awareness of their strengths and shortcomings as leaders. Best practices literature has documented the importance of

creating multiple avenues for feedback and using data to create constructive development plans (Creamer & Janosik, 2003). Instrumented leadership assessment can serve an important role in this process, provided instrumentation is reliable, measures valid constructs, and is appropriate for the population (Creamer & Janosik, 2003).

The synergistic supervision model articulated the importance of dynamically balancing complementary but opposing demands (Armenio & Winston, 2001; Winston & Creamer, 1997; 1998; Winston & Hirt, 2003; Winston, Ullum, & Werring, 1984). Validated instruments such as the SLPI measure key aspects of leadership, but leadership is a broad phenomenon and much is missed. Existing student leadership instruments cannot measure a leader's ability to dynamically balance complementary but opposing strengths. The LVI-S seeks to remedy this limitation by providing insight into housing and residence life supervisors' leadership versatility.

There are many questions to consider about whether leadership versatility and the LVI-S are valid in a housing and residence life context. The following discussion reveals interesting and relevant results about measuring versatility in this setting. It considers whether the LVI-S effectively replicates or approximates the psychometric performance of the executive LVI, whether or not versatility is a valid measurement construct for the housing and residence life context, and whether the LVI-S might inform future training, research, and development of housing and residence life leaders.

Research Question 1

Validity may be considered the degree to which an instrument measures what it intends to measure (Kaiser & Kaplan, 2006). Since the underlying design of the LVI-S

was based on the executive LVI, a primary consideration was whether the new instrument could reasonably replicate or approximate the factorial structure and characteristics of the original instrument. Research question 1 assessed whether the LVI-S maintained the properties necessary to measure versatility across the Forceful and Enabling leadership dimensions in a campus context. Although LVI-S dimensions and sub-dimensions were conceptually similar to the original, item language was modified or redesigned because the lexicon of business professionals differs from that of residence-hall staff. The LVI-S was expected to demonstrate the following criteria: 1) Forceful and Enabling leadership are distinct factors, not opposite ends of a single continuum; 2) Both dimensions should yield three distinct subcomponents; and 3) the Forceful and Enabling dimensions should be inversely related to one another.

A series of principal components analyses (PCA) demonstrated the LVI-S was able to approximate a great portion of the original instrument's factorial structure, though each dimension exhibited unique strengths and weaknesses. The PCA of all 24 items in the Forceful and Enabling dimensions showed 22 out of 24 loaded on either a Forceful component or an Enabling component. The two exceptions exhibited strong negative loadings on their complementary opposite component. Item *Fd_8*, from the *Declares* sub-dimension loaded negatively (-.61) on the same component as items from the *Listens* sub-dimension. Item *Es_24*, from the *Supports* sub-dimension, loaded negatively (-.59) on the same component as three out of four *Pushes* items. This 24-item analysis provided evidence the Forceful and Enabling dimensions of the LVI-S were distinct constructs.

To ease interpretation, the 24-item PCA was segmented into two separate, 12-item PCAs. In these analyses the Forceful dimension yielded four components and the Enabling dimension yielded three. The Forceful dimension provided the most organized item-loading characteristics. The Forceful component matrix showed the majority of items loading on their intended sub-dimension. Cross-loading was limited to three items. In two of the three instances, items loaded on their intended component but then cross-loaded on an “extra” component that emerged during the extraction.

This “extra” component, identified as component 3, appeared to represent a latent variable occupying factor space between *Takes Charge*, *Declares*, and *Pushes*. Language for these items exhibited a bias toward overt action: “gives clear direction,” “decisive,” “*Declares*,” and “holds people accountable.” Item Ftc_3, “Gives clear direction—tells people what to do” was intended to measure the *Takes Charge* construct. Item Fd_5, “Decisive—makes up his/her mind quickly” was expected to load with other items from the *Declares* sub-dimension. Fp_12, “Holds people accountable—is firm when others do not meet his/her standards” loaded almost equally on component 3 and component 1. Fd_6, “Lets people know clearly where he/she stands on issues—*Declares* him-/herself” loaded strongly on component 2, but cross-loaded on component 3. A common characteristic of these items was assertiveness—as if items from all three sub-dimensions were rated on an assertiveness continuum and the strongest sounding item from each was placed in a common category.

Psychological measurement literature refers to a phenomenon known as the difficulty factor artifact (Nunnally & Bernstein, 1994). Items exhibiting similar levels of

difficulty can group together if respondents tend to rate them in a similarly high or similarly low manner. Some items on the LVI and LVI-S are intentionally provocative, attempting to draw out salient aspects of overdoing behavior. It is possible the housing and residence life population exhibited collective sensitivity to more assertive aspects of Forceful leadership. Although this explanation may be theoretically possible, it would be preferable to conduct further analyses before resorting to the difficulty factor explanation. To better determine the components comprising the Forceful dimension, a bi-factor hierarchical model might be used (Reise, Morizot, & Hays, 2007). This process allows items to load on a general trait that is assumed to underlie all the items, such as the common aspects of the Forceful dimension, but it also permits items to load on one or more group factors. In the case of the Forceful sub-dimension, the bi-factor model could be constructed to allow items to load on the general factor (Forceful leadership) as well as on group factors (e.g., the three sub-dimensions). Such an analysis is beyond the scope of the current study but will be considered for future research with the LVI-S.

The Enabling dimension PCA exhibited different characteristics. As anticipated, the extraction produced three components but item loadings were not organized as neatly as in the Forceful dimension. Component 2 provided the most distinct representation of a sub-dimension; all four *Empowers Others* items loaded on this component. The other two components were less clear.

Seven items loaded on component 1, indicating it was comprised of more than one sub-dimension. Of the seven items, three were contributed by *Listens*, three were contributed by *Supports*, and one cross-loaded from component 2. The *Supports* items

loading on component 1 were fairly generic statements: Es_22, “Friendly—is nice and treats other people with courtesy”; Es_23, “Sensitive—careful not to hurt the other person’s feelings”; Es_24, “Cuts people slack—is understanding when they are not able to do their duties and responsibilities.” Overall, it appeared the more general concept of supporting others was analogous to the act of listening. Component 1 appeared to indicate the majority of *Listens* and *Supports* items did not represent two distinctly different sub-dimensions.

There was a third component, however, that could be interpreted as evidence of a *Supports* component. Since component 3 contained only two items, it should not be interpreted as full representation of a sub-dimension. Even so, these items addressed aspects of supportive behavior such as respect, openness, and validation of others. The items were: El_20, “Makes it safe to challenge or critique his/her thinking—welcomes dialogue and debate” and Es_21 “Shows appreciation—goes out of his/her way to make others feel good about their contribution.” These items could provide a foundation for development of a more distinct *Supports* sub-dimension. In all, five out of six sub-dimensions were fully represented in the LVI-S.

Although the LVI-S largely replicated the factor structure of the LVI, it needed to also demonstrate an inverse relationship between the Forceful and Enabling dimensions. A correlation analysis evidenced a significant, negative relationship between the Forceful and Enabling dimensions ($-.47, p < .001$). In summary, these PCA and correlation analyses indicated the LVI-S successfully approximated or replicated key criteria

necessary for measuring Forceful and Enabling leadership versatility in a housing and residence-life context.

Research Question 2

The next research question explored the predictive validity of the LVI-S sub-dimensions. Did the sub-dimensions relate to other non-test variables in a manner consistent with theory? To test the hypotheses, sub-dimension scores served as independent variables and effectiveness ratings provided by RAs served as the dependent variable. The effectiveness variable was a composite of three effectiveness items ($\alpha = .853$) derived from Tsui's (1984) research into reputational effectiveness. Hypothesis 2a proposed RA-reported sub-dimension scores would significantly predict effectiveness ratings. Hypothesis 2b proposed that HD-reported sub-dimension scores would not significantly predict effectiveness ratings. Regression analyses were used to test both hypotheses.

Hypotheses 2a tested how well observer-reported sub-dimension scores predicted HD effectiveness. Fundamentally speaking, did sub-dimensions measure behaviors relevant to effective leadership? The findings supported hypothesis 2a. Observer sub-dimension ratings were strongly correlated with effectiveness ($R = .75$; $R^2 = .56$, indicating sub-dimensions explained more than 50% of the variance in effectiveness. Adding credence to the Forceful / Enabling duality model, sub-dimensions on both sides of the duality were significant predictors. In order of strength, the predictors were *Listens* (Beta = .78, $t = 5.01$, $p < .01$), *Takes Charge* (Beta = .47, $t = 2.7$, $p < .05$), *Pushes* (Beta =

-.43, $t = 2.98$, $p < .01$), *Empowers Others* (Beta = -.29, $t = 2.18$, $p < .05$). The *Supports* and *Declares* sub-dimensions were not significant predictors in the model.

Listening was most strongly associated with effectiveness. This sub-dimension measured inclusive leadership behavior, feedback seeking behavior, openness to influence, and openness to competing points of view. The Listening coefficient was nearly twice that of the next closest predictor, an indication that RAs believe these behaviors are closely aligned with being an effective HD. *Takes Charge* was the second highest predictor. RAs believed effective hall directors provide clear direction, are perceived to be in control, and are willing to get involved should problems arise. The *Pushes* sub-dimension was nearly as strong a predictor as *Takes Charge*, but it produced a negative coefficient, indicating *Pushes* was associated with lower effectiveness scores. Behaviors included pushing for high performance, expecting a lot, and providing direct feedback when expectations have not been met. *Empowers Others* yielded a negative coefficient as well. RAs did not associate effective leadership with HDs who are perceived to be hands-off, who give people room to do their jobs, and who trust staff to handle problems as they arise. It may sound counter-intuitive to correlate these behaviors with ineffective leadership. The question is not whether these behaviors should or should not be done, but rather in what situation or to what degree they are done. A more in-depth discussion of situation and degree will be addressed in the discussion of research question 6.

The regression coefficient in the preceding analysis was quite high ($R = .75$; $R^2 = .56$) but it should be interpreted with care. The analysis used to test the hypothesis

incorporated independent and dependent variables provided by a common source, making the results susceptible to mono-source bias. Podsakoff and Organ (1986) cautioned that leadership ratings collected from one source and effectiveness ratings collected from the same source may produce correlations that appear more substantial than they really are. Many well-known leadership instruments incorporate an effectiveness scale within the instrument (e.g., Multifactor Leadership Questionnaire), so the practice is not without precedent. Should the regression from hypothesis 2a be invalidated due to mono-source bias? If results were closer to the significance threshold, then a case could be made that they should be. Given the strength of the correlation, however, correcting for mono-source bias would likely still yield a significant value. Podsakoff and Organ (1986) did not suggest mono-source correlations were complete artifacts but rather indicated that the strength of the correlation could appear exaggerated. Thus, the overall strength of the regression coefficient should be tempered by this understanding.

Hypothesis 2b explored whether HD self-report scores would predict effectiveness. Whereas the first hypothesis tested the predictive validity of LVI-S sub-dimensions, the second tested the validity of self-report data. A few studies have explored observer versus self-ratings validity in residence hall settings. In a study using the Multifactor Leadership Questionnaire (MLQ), Komives (1999a) identified discrepancies between HDs' and RAs' estimates of the leader's level of positive engagement, disengaged leadership characteristics, and non-leadership behaviors. Komives (1991b) also reported discrepancies between HD estimates and RA reported levels of satisfaction and motivation. Turrentine (2001) used a self-authored Leadership Skills Assessment to

compare residents' and peers' perceptions of leadership behaviors. The overall analysis indicated a significant difference between self-reports and observer reports ($t=3.163$, $df=107$, $p = .002$). Curiously, Turrentine's commentary ignored the findings, "This study began with the observation that student affairs scholarship and practice rely on student self-reports. The results of this study, if confirmed in future research, provide a basis for confidence in students' accounts of their own behaviors" (p. 371). Turrentine (2001) has been cited as empirical support for self-report methodology, "... a study on self-and peer reported leadership behaviors and the quality of those behaviors found self-reports of leadership to be generally accurate" (Dugan, Garland, Jacoby, & Gasiorsky, 2008, p. 288). Self-report methods are an important and prevalent aspect of student leadership research—both in housing and residence-life as well as broader campus settings. Few campus-based studies have explored the validity of self-versus-other data and the existing research indicates mixed findings that are not corroborated in the larger body of leadership research.

The regression analysis for 2b was conducted in the same manner as that in 2a. Sub-dimension scores were used to predict RA-provided effectiveness scores. In this analysis, however, sub-dimension scores were comprised of self-reported data from HDs. Hypothesis 2b predicted that HDs self-reported scores would not significantly predict effectiveness. The regression analysis was not significant, indicating support for the proposed hypothesis. Self-reported data did not predict the criterion variable as effectively as the observer-reported data.

A large body of multi-rater research has indicated self-reported data is not as reliable or valid as observer-reported data (Atwater & Yammarino, 1992; Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988; Komives, 1999b, Posner, 2005). Leadership is a social behavior (Denison, Hoojberg, & Quinn, 1995). Leaders engage in observable actions that represent effective or ineffective leadership. External reports of these actions are more reliable and valid because they are based on observation, whereas self-reported data is biased by individuals' thoughts, feelings, rationalizations, and intentions (Atwater & Yammarino, 1992; Conway & Huffcutt, 1997; Harris, & Schaubroeck, 1988). On the whole, leaders are not considered to be reliable evaluators of their own leadership.

Even so, a secondary analysis was conducted to further test the validity of self-reported data. This follow-up analysis was more conservative than the original. Instead of using self-reported behaviors to predict a secondary construct (effectiveness), the follow-up test sought to determine whether self-rated sub-dimension scores were able to predict observer scores for the same sub-dimension. For example, an HD's (self-reported) *Takes Charge* score would serve as the independent variable and the RAs' average (observer-reported) *Takes Charge* score would serve as the dependent variable. Should the regression prove significant, it *Supports* the validity of self-report data. If the regression proves insignificant, then it provides additional evidence that the validity of self-reported scores should be called into question.

When the six sub-dimensions were tested, only one sub-dimension, *Supports*, yielded a significant result. Of the other five sub-dimensions, *Listens* provided the most compelling evidence against self-report validity. HD self-ratings explained only 3% of

the variance in RA ratings of *Listens* ($R = .16$, $R^2 = .03$). Comparing this finding with results from the initial regression is instructive. When RA-reported sub-dimension scores were used to predict effectiveness, *Listens* ($\beta = .78$) was clearly the strongest predictor. RAs who rated HDs as good listeners also considered those HDs to be effective leaders. A comparison between self- and observer-scores, however, indicated HDs were thoroughly unable to discern whether RAs had favorable or unfavorable perceptions of their listening performance. To further illustrate this point, when HD self-scores were used to predict effectiveness, Listening was the strongest predictor (see Table 17). But in the HD regression, *Listens* was inversely related to effectiveness ($\beta = -.31$). This discrepancy does not make a good case for validating self-reports of performance.

The findings from hypothesis 2b add to the evidence against the validity self-ratings of leadership behavior. These results extend findings reported by Komives (1991b). Although the evidence corroborates *statistical* findings of Turrentine's (2001) study, it completely opposes the author's speculation that self-reports of leadership behavior can serve as a valid source of behavioral data.

Creamer and Janosik (2003), Brown (1988), and Grote (1996) referenced the importance of behavior-based multi-rater measurement for the development of student-affairs professionals—although the difficulty in finding valid and reliable instrumentation was noted as well. In housing and residence-life, multi-source feedback has been recommended for the development and supervision of RAs and HDs (Upcraft & Pilato, 1982; Winston & Fitch, 1993). A handful of multi-rater leadership studies have been published in college student development journals over several decades, (Adams &

Keim, 2002; Komives, 1991a; 1991b; 1991c, Posner, 2005; Posner & Brodsky, 1992; 1993; 1994; Posner & Rosenberger, 1997); but limited research has been conducted outside of Posner's notable efforts with the SLPI.

The broader body of student leadership development research makes little reference to the value of instrumented, multi-rater feedback. Pascarella and Terenzini's (2005) comprehensive review of student development research observed, "the post 1990 [leadership] research has largely ignored interpersonal relations in favor of examining the influences of institutional characteristics on students' *self-reported* abilities in areas related to leadership" (p. 236, italics added). Self-report instrumentation and methods can serve a valuable purpose, but where leadership behaviors are concerned, other forms of assessment should complement self-rater methods. Ignoring multi-rater research methods and observer-based feedback undermines the quality and richness of student leadership research.

Research Question 3

Research question 3 explored the convergent construct validity of the LVI-S. Constructs measured by new instruments can be correlated to those measured by known and validated instruments to test for convergent validity. Several well-known instruments measure constructs similar to those in the LVI-S [e.g., Stogdill and Coons' (1957) LBDQ], however none of these was designed with students or campus settings in mind. The Leadership Practices Inventory—Student Version (Kouzes & Posner, 2003) is an established, student-focused multi-rater instrument. It is the only one validated on a student population within a campus context (Schwartz & Gimbel, 2000). Since the

purpose of the study was to develop and validate a multi-rater instrument for college students, selecting a comparison instrument that had been tested within the target population was a priority. The SLPI measures transformational leadership constructs (Sashkin, 2004). Transformational constructs generally focus on egalitarian aspects of leading: initiating change for the greater good, building consensus toward shared vision, and generating commitment to achieve meaningful objectives (Bass, 1986; Burns, 1978; Kouzes & Posner, 2002). Because of the strong emphasis on relational leadership, the Enabling sub-dimensions were expected to correlate with SLPI scales more than the Forceful sub-dimensions. This imbalance is not ideal because it could limit the ability of the SLPI to validate both dimensions of the LVI-S, however the SLPI measures some constructs that similar to Forceful concepts—such as Challenge the Process. The presence of these constructs mitigated some of the risk of selecting a transformational instrument.

The SLPI is composed of 5 scales: Models the Way, Inspires a Shared Vision, Challenges the Process, Enables Others to Act, and Encourages the Heart. The first hypothesis proposed the Enabling sub-dimensions of the LVI-S would correlate significantly with Enabling Others to Act, Inspiring a Shared Vision, and Encouraging the Heart scales on the SLPI. *Empowers Others* correlated significantly with one SLPI scale, Enables Others to Act ($r = .17, p < .01$). The *Listens* sub-dimension correlated significantly with all three scales (Enabling Others, $.54, p < .01$; Vision, $.34, p < .01$; Encouraging, $.37, p < .01$). Similarly, the *Supports* sub-dimension produced significant

correlations with all three scales (Enabling Others, $.43, p < .01$; Vision, $.28, p < .01$; Encouraging, $.38, p < .01$).

The Enabling dimension's egalitarian leadership behaviors align with transformational constructs in some obvious and not-so-obvious ways. Unexpectedly, strong relationships also were found for Models the Way and Challenge the Process. Both scales correlated with *Listens* (Models, $r = .42, p < .01$; Challenge, $r = .33, p < .01$) and *Supports* (Models, $r = .32, p < .01$; Challenge, $r = .22, p < .01$), adding further support for hypothesis 3. Enabling sub-dimensions measured actions leaders take to “enable” others to step up and lead.

The Forceful dimension is the complementary opposite of the Enabling dimension. Theoretically speaking, Forceful sub-dimensions should exhibit different relationships with the SLPI than the Enabling sub-dimensions. Differences could be evidenced in terms of frequency (fewer significant correlations), degree (lower r), and direction (negative correlations). Hypothesis 3b suggested one SLPI scale, Challenge the Process, would significantly correlate with the three Forceful sub-dimensions. Limited support was found for this hypothesis. The *Takes Charge* sub-dimension was significantly correlated with Challenges the Process ($r = .33; p < .01$), but *Declares* and *Pushes* were not. Three out of four *Takes Charge* items correlated significantly with almost every item in the Challenges scale—the remaining *Takes Charge* item exhibited significant correlations, but to only two out of six Challenges items.

Overall, the Forceful dimension did not exhibit as many significant relationships as the Enabling dimension. *Declares* and *Pushes* did not provide any positive

relationships of statistical significance. The lack of correlation indicates the sub-dimensions measured different constructs than the SLPI scales. This finding does not indicate the Forceful sub-dimensions measured their intended constructs, but it does distinguish them from the SLPI constructs and their closely related Enabling sub-dimensions.

Several unexpected and informative relationships emerged in the Forceful dimension. First, *Models the Way* correlated significantly with *Takes Charge*. Item Ftc_3, “Gives clear direction” and item Ftc_4, “Steps in—gets actively involved when problems arise” produced the majority of significant correlation. This relationship should have been anticipated because leaders “model” effective leadership when they provide clear direction, establish expectations, and exude a sense of competence and control. Second, two of the most interesting results were cases of discriminant validity. The *Enables Others to Act* scale from the SLPI was strongly correlated with all three Enabling sub-dimensions. On the Forceful side of the duality, however, *Declares* and *Pushes* exhibited significant negative correlations with *Enables Others to Act*: *Declares* ($r = -.21, p < .01$) and *Pushes* ($r = -.21, p < .01$). These inverse relationships provided further evidence of the polarity effect in the LVI-S: Forceful and Enabling constructs should negatively correlate with one another. Although *Declares* and *Pushes* were not validated through a positive correlation, the fact that they correlated negatively with an opposing construct provided evidence of discriminant validity. In summary, the LVI-S and the SLPI related to each other in expected, as well as unexpected but theoretically

understandable ways. The significant correlations found in the analyses provided evidence of convergent and discriminate validity for the new instrument.

Research Question 4

The next research question explored whether Hall Directors over-estimated, under-estimated, or congruently estimated their self-ratings in comparison to observers. Assigning rater categories and discerning the implications of these categories is a well-documented practice within multi-rater research (Atwater & Yammarino, 1992; Beehr, Ivanitskaya, Hansen, Erofeev, and Gudnaowski, 2001; Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988). Literature indicates individuals are poor estimators of their own leadership (Beehr, et al., 2001; Conway & Huffcutt, 1997; Harris & Schaubroeck, 1988). The findings from hypothesis 2b indicated similar results for Hall Directors. First, HD self-scores were unable to predict RA ratings of effectiveness. Second, a more conservative test examined how well HDs' self-rated sub-dimension scores could predict observer scores. In five out of six cases, HD self-scores were unable to predict observer scores—though the *Supports* sub-dimension provided a notable exception.

It is not uncommon for self-ratings to be inflated in comparison to observer ratings (Atwater & Yammarino, 1992; Beehr, et al., 2001). Many theories have been offered to explain why, including egocentric bias (Conway & Huffcutt, 1997), ambiguity of the rating context (Harris & Schaubroeck, 1988), level of self-esteem (Conway & Huffcutt, 1997; Atwater & Yammarino, 1992), and lack of knowledge or insight for effective self-other comparisons (Atwater & Yammarino, 1992; Kruger & Dunning,

1999). Although it was clear HD self-ratings were unable to predict observer ratings, was this because they were over-estimating or under-estimating?

Rater categories were defined for under-estimating, over-estimating, and for those in agreement. Categories were assigned separately for the Forceful dimension and for the Enabling dimension and a crosstab analysis summarized the results (see Table 22). The proposed hypothesis stated that a majority of HDs would be categorized as over-estimators, but the results indicated otherwise. The results indicated that 21 HDs (42.9%) in the Forceful dimension were categorized as In Agreement with their observers, 12 (24.5%) were Under-estimators, and 16 (32.7%) were Over-estimators. A similar pattern emerged in the Enabling dimension: 22 (44.9%) HDs were In Agreement, 13 (26.5%) were Under-estimators, and 14 (28.6%) were Over-estimators.

During the course of this research, several important characteristics about rater agreement terminology emerged. An important distinction should be made between rating inflation and over-estimation. Rating inflation refers to the general tendency for self-ratings to be somewhat higher than observer ratings. Ratings inflation can occur on a continuum (Atwater & Yammarino, 1992). The *over-estimation agreement category*, however, is a function of ratings inflation with respect to a distribution of difference scores within the sample population (Atwater & Yammarino, 1992). Thus, over-estimators exhibit a significantly higher degree of inflation than the average case. At the time Hypothesis 4 was conceived, the author was unclear about the distinction between rating inflation and over-estimation. Hypothesis 4 stated most HDs would over-estimate

their leadership ratings in comparison to those provided by RAs. As the analysis progressed, it became clear the hypothesis was a poor fit to the data.

Agreement groups were calculated from a distribution of difference scores. Should the distribution be normal, a majority of over-raters could not over-estimate because the parameters were relative to the observed distribution. Consider the Forceful dimension difference score statistics presented in Table 21 ($m = -.291$; $\frac{1}{2} sd = .228$; Under-estimator Lower Bound $< -.519$; Over-estimator Upper Bound $> .063$). The agreement group parameters were based off of the mean difference of all difference scores ($m = -.291$). An Upper Bound was calculated by adding $\frac{1}{2} sd (.228)$ to the mean ($-.291$), yielding an upper parameter (Upper Bound $> -.063$). Cases with difference scores that exceeded this parameter were Over-estimators. The lower parameter was calculated by subtracting $\frac{1}{2} sd (.228)$ from the mean ($-.291$), producing the lower parameter (Lower Bound $< -.519$). Cases with difference scores less than the lower parameter were under-estimators. Cases that fell in between the two parameters were considered in agreement.

Dividing a normal distribution in the manner prescribed by Atwater and Yammarino (1992) will place a majority of raters in the in agreement category. The remaining difference scores will be more extreme, albeit less frequent, and more or less evenly distributed into over- and under-estimator categories. Results in Table 22 indicate the Forceful and Enabling difference score results followed the normal trend. For Hypothesis 4 to fit the data, a distribution would have needed to be bimodal, with greater frequency on the over-estimator side, or it would have to have been heavily skewed toward over-estimation. There was no prior research to suggest either of these

distributions was likely, therefore the hypothesis was flawed. This flaw occurred because during the conceptual phase of this study, the term “rater inflation” was considered analogous to being an over-estimator. The evidence proved this is not the case, at least with respect to the methods used in this analysis.

Research Question 5

Understanding the differences and implications of rater agreement categories is important with respect to leadership research and the development of leaders. Atwater and Yammarino (1992) reported that self-awareness moderated the relationship between leader behavior and correlations with predictor variables. Significant findings varied in relation to rater agreement category and rating source (superior versus subordinate), therefore research should not only distinguish self-aware raters, but also differentiate between under-estimators and over-estimators. If over- and under-estimators are not distinguished, moderating effects from these two types of non-agreement could alter or cancel out significant correlations (Atwater & Yammarino, 1992).

In the Atwater and Yammarino (1992) study, leader behavior was positively related to predictor variables (e.g., leader performance) for individuals in the under-estimator and self-aware categories. Under-estimators and self-aware individuals appeared to learn from prior experiences and use feedback from these experiences to modify their behavior (Atwater & Yammarino, 1992). Kruger and Dunning (1999) reported similar results in a series of separate non-leadership studies.

Over-estimators’ experiences and abilities were negatively related to leader behavior ratings (Atwater & Yammarino, 1992). Over-estimators did not appear to

effectively learn from experience. The authors speculated that learning was limited by a self-serving perceptual bias. This bias hindered over-estimators' capacity to recognize constructive feedback and transfer it into behavioral change (Atwater & Yammarino, 1992). Kruger and Dunning (1999) reported similar limitations for over-estimators, but they believed over-estimation may result from lack of knowledge or insight—the individual may be “unskilled and unaware” (p. 1121). Additional studies indicated over-estimators exhibit lower performance on criterion variables (Beehr, et al., 2001; Komives, 1991b; Kruger & Dunning, 1999) and are more susceptible to career derailment (Lombardo & Eichinger, 2000). Some have reported over-estimation might be a stable characteristic that continues throughout life (Nelson & Campbell, 1993).

What implications do agreement categories have with respect to leadership effectiveness in the residence hall setting? Hypothesis 5a suggested there would be a statistically significant difference between the mean effectiveness scores of over-estimators versus under-estimators. Hypotheses 5b proposed over-estimators were expected to receive lower effective scores than under-estimators. Analyses were conducted separately for the Forceful and Enabling dimensions.

A one-way ANOVA on the Forceful dimension compared effectiveness scores for the self-aware, under-estimator, and over-estimator agreement categories. The initial ANOVA did not identify any significant mean differences for effectiveness. Follow-up analyses on the sub-dimensions were conducted to discern whether differences were present at a more discrete level of data. One sub-dimension, *Takes Charge*, produced a significant result ($F = 5.9, p < .01$). The post-hoc analysis determined the difference was

due to a higher mean score for the self-aware group. Although this result did not support the proposed hypothesis, it did support Atwater and Yammarino's (1992) findings regarding self-awareness and positive predictors of leadership.

Results for the Enabling dimension provided support for hypotheses 5a and 5b. The dimension-level ANOVA was significant ($F = 6.88, p < .01$). A post-hoc analysis reported the mean effectiveness score for over-raters ($m = 15.03$) was significantly lower than that for under-raters ($m = 17.46$) and the self-aware group ($m = 17.07$). In light of previous research on transformational leadership and RA perceptions of effectiveness, it makes sense that effectiveness ratings could be highly correlated to enabling behaviors. Transformational leadership is highly related to RAs' perceptions of satisfaction (Komives, 1999b; Paladino, Murray, Newgent & Gohn, 2005) and leader effectiveness (Komives, 1999; Posner & Brodsky 1993). The behaviors measured by the Enabling sub-dimension are highly correlated with transformational leadership. When RAs perceive strengths or shortcoming in these behaviors, effectiveness ratings are likely to reflect them.

The support for hypothesis 5a and 5b add evidence to a well-documented outcome: transformational-type behaviors can have significant impact on RA effectiveness ratings. Additionally, it extends findings from Atwater and Yammarino (1992) and Kruger & Dunning (1999) into the realm of residence hall leadership. Further examination of agreement categories, sub-dimension mean scores, and effectiveness ratings provides insight and clarifies the implications these results may hold for residence life leadership development.

Parsing data can reveal hidden relationships. Atwater and Yammarino (1992) stated that self-awareness moderated the relationship between leader behavior and correlations with predictor variables. Significant findings varied in relation to rater agreement category and rating source (superior versus subordinate). If over- and under-estimators are not distinguished, moderating effects from these two types of non-agreement could alter or cancel out significant correlations (Atwater & Yammarino, 1992). The present study provides evidence of these effects.

An example of this masking characteristic may be seen in the *Empowers Others* sub-dimension in Table 26. In the sub-dimension, two out of three agreement group correlations between *Empowers Others* and behavior were insignificant, but the Aware category was an exception (*Empowers Others, Aware, r = .55, p < .05*). When agreement categories were excluded from the analysis, this significant correlation vanished (*Empowers Others, Aware, r = .08*).

A comparison of the self-aware and over-estimator categories shows very few mean differences in self-ratings. On a 9-point scale, the differences only ranged from .04 to .2. Since self-scored mean differences were minimal, the difference between the groups must be due to observer scores. The self-aware group mean for *Listens* was 5.11 and the over-raters' mean was 5.07. RAs rated the self-aware leaders an average of 4.8, but the over-estimators received just a 4.09, nearly a one-point lower difference. This ratings gap highlights a potentially serious blind spot for the over-estimators. Listening is highly related to effectiveness (*Aware, r = .65, p < .01; Over, r = .80, p < .01*), and at

least some of those in the over-estimator group think they are much better listeners than they are.

If housing and residence life directors were to use the LVI-S for staff development, they could use this tool to target specific areas for growth. Multi-rater instrumentation is a viable form of feedback because recipients have appreciated the specific and candid results these instruments can deliver—provided it is delivered by someone trained in fundamentals of effective feedback (Brown, 1996; Fleenor & Leslie, 1998; Winston & Fitch, 1993). Were an HD to over-estimate listening skills, this could be an indication the leader has a blind spot in a very important domain. Developmental training, coaching, and supervision can collectively serve to support this individual's growth.

Training programs designed around multi-rater protocols have demonstrated success for many decades (Eichinger & Lombardo, 2003; Leslie & Fleenor, 1998; McCauley & Van Velsor, 2004; Parry, 2005). Proper use of multi-rater assessment methods in campus settings can enhance the quality of research produced. When combined with effective feedback delivery, multi-rater assessment also has the potential to improve leadership outcomes (Atwater & Yammarino, 1992; Komives, 1991b; Posner, 2004; Posner & Brodsky, 1993).

Research Question 6

The predictive validity of LVI-S dimensions and sub-dimensions was explored in question 2 through a simultaneous regression. This regression is considered an additive model approach because effectiveness was predicted by the collective power of each sub-

dimension. The results from research question 2 were encouraging, but the additive model does not test the predictive validity of versatility. The final research question examined this issue.

The root of the word versatility meant “to turn around” or “to pivot” (Kaplan, 1996, p.1). Kaplan conceptualized versatility as the ability to pivot from one style or approach to another depending on the needs of a situation. Kaplan and Kaiser (2003; 2006) presented leadership versatility as a dynamic relationship between seemingly opposite virtues. In theory, the LVI-S is designed to measure relationships across the Forceful / Enabling duality. Versatility is calculated through joint-standing variables that consider the interaction (or joint-standing) of complementary, but opposed behaviors. A more specific explanation of these calculations is articulated in Chapter 3 and in Appendix X.

The majority of LVI-based research has been conducted with corporate executives. Kaplan and Kaiser (2006) reported the relationship (R) between versatility and effectiveness ranged between .60 and .79 across 5 samples of executives ($n = 450$). The average multiple correlation ($R = .71, R^2 = .50$) indicates versatility accounted for half of what it meant to be regarded as an effective leader in the corporate suite.

Is versatility relevant in a different leadership context? Interestingly, the synergistic supervision model (Winston, Ullom, & Werring, 1984) appears to share much in common with Kaplan and Kaiser’s versatile leader model. Effective residence hall supervision requires a dynamic balance of leadership qualities (Winston & Fitch, 1993). Armenio and Winston (2001) qualitatively examined behaviors, values, and attitudes that

characterized high quality supervisors. Among other characteristics, quality supervision included forceful behaviors such as setting the context, giving direction, and motivating; conversely, these were synergistically balanced with observing, listening, and caring behaviors (Arminio & Winston, 2001). Creamer and Janosik (2003) emphasized the importance of instrumented assessment and recommended a range of performance appraisal tools for use in student affairs. None was capable of measuring the dynamic relationships necessary for measuring versatility or synergistic supervision.

In every case but one, versatility out-predicted the additive model (Table 27). Two findings warrant particular note. First, it is instructive to note the difference in the predictions for the *Takes Charge / Empowers Others* sub-dimensions. The versatility model more than doubled the explained variance of the additive model (versatility R Squared = .30; Additive R Squared = .13). This sub-dimension pairing carried the strongest β in the Forceful / Enabling versatility regression with all three sub-dimension pairs.

A comparison between the LVI-S and the SLPI provided several points of interest. The transformational scales from the SLPI were strongly correlated with effectiveness. In comparison to the LVI-S, the predictive validity of the SLPI was more consistent across each of its scales. The SLPI demonstrated excellent reliability characteristics and its psychometric stability surely aided these predictive characteristics. The LVI-S will benefit from refinements that tighten the inter-correlation of sub-dimension items. Such development will enhance inter-rater reliability, improve psychometric stability, and enhance its predictive ability. Table 14 shows that although

the LVI-S performed sufficiently, further development is warranted. In spite of early-stage growing pains, the predictive power of versatility was noteworthy. In its current guise, the LVI-S explained 36% of what it meant to be an effective Hall Director. The versatility prediction model out performed the additive model using the LVI-S scales and slightly out performed the 5 SLPI scales as well (LVI-S versatility model, $R = .60$, Adj. $R^2 = .31$, $F = 7.72$, $p < .01$; LVI-S additive model $R = .55$, Adj. $R^2 = .20$, $F = 2.87$, $p < .05$; SLPI additive model $R = .58$, Adj. $R^2 = .25$, $F = 4.05$, $p < .01$) This result validated the importance of versatility in a housing and residence life supervision context. Given the high degree of conceptual similarity between Kaplan and Kaiser's (2006) versatile leader model and the synergistic supervision model, this finding provides compelling evidence for future research and training into these areas.

Until this study, Winston and colleagues could intuitively discern the importance of synergistic supervision and study it through qualitative methods, but it was difficult to measure quantitatively. Creamer and Janosik (2003) observed that scales can be difficult and time-consuming to develop, but if resources are applied wisely, the feedback they generate can be invaluable. The LVI-S offers an efficient and effective measurement method and its behavior-based nature lends makes it easy to apply as a supervision and feedback tool in residence life and student affairs settings.

Limitations

The presented study is limited by the following constraints. The sample was a convenience sample comprised primarily of institutions in the Southeastern United States. A variety of institutions sampled (an HBCU, two small private colleges, and four

universities), but caution should be used before generalizing these conclusions to regions and institutions outside the sample parameters. It is possible some respondents could have provided ratings that were overly harsh or overly positive; the sampling method could not control for this possibility. The institutions surveyed varied greatly in size and in the number of RAs supervised by the HD. Institutions with lower HD to RA ratios were not as well represented within the sample. Therefore results may generalize more effectively to larger institutional contexts.

The present study used effectiveness ratings as a criterion variable. Observers (RAs) provided Hall Director (HD) effectiveness ratings used in these analyses. Leaders' behavioral ratings and effectiveness ratings were provided by a common source: RA observers. This measurement method can produce common source bias and some evidence of this bias was identified. Where possible, this issue was addressed within the study. When it was not possible to avoid, these limitations were noted in the results and subsequent discussion of results. Future research efforts may want to consider methods for obtaining effectiveness ratings from independent sources (e.g., supervisor effectiveness ratings or objective performance-based metrics).

Implications for College / University Counseling

The psychometric support for this measure, as well as the findings of these research questions, have implications for counselors and higher education personnel charged with developing leaders on college campuses.

Leadership development and counseling share many processes and underlying intentions to encourage growth and development. In cognitive-behavior counseling and

solution-focused therapy, development plans are founded on explicitly stated goals. Interventions are constructed to help a client practice new behaviors and, ideally, learn new ways of thinking that reinforce behavior change. In many respects, leadership development follows a similar plan. In professional leadership development, leadership coaches help clients identify areas of growth and help them make these changes. In housing and residence life, developmental supervision can provide guidance for growth and professional development. The synergistic supervision model prescribes this approach. Experts in both professional and residence life domains recommend using objective, candid, and fair assessments to provide the foundation for developmental dialogue.

The results of this study provide empirical support for the relevance of leadership versatility in housing and residence life settings. They also validate precepts underlying the synergistic supervision model. The LVI-S provides a quantitative, behavior-based method for measuring synergistic supervision. This function serves dual purposes.

First, the LVI-S may be used as an assessment tool for developing professional and paraprofessional leaders within the housing and residence life context. LVI-S feedback can help supervisors and supervisees co-construct development plans that address over-used and under-used behaviors. Thus, the LVI-S may be helpful for the development and supervision of individual leaders.

Additionally, the LVI-S may be used as a research instrument for organizational development. The current collection of campus-based leadership assessments is sparse, especially where behavior-based multi-rater instruments are concerned. Most multi-rater

instruments used in housing and residence life settings measure outcome variables such as residence hall climate and resident satisfaction, they are not validated to measure leadership behaviors. These outcome variables, however, may be mediated or moderated by the quality of leadership provided by the supervisors and administrators of the residence hall. The LVI-S may be used to test for these mediating or moderating relationships.

The LVI-S may be especially well suited for such research because it measures the Forceful and Enabling dimensions of leadership. Conceptually, the Forceful dimension aligns closely with the administrative, institutional responsibilities identified in the synergistic supervision model. The Enabling dimension aligns with more supportive, individualized aspects of the model. Furthermore, the LVI-S measures over-use of behaviors. Since the synergistic supervision model is predicated on balancing between complementary demands of supervision, it makes sense that it is measured through an instrument that distinguishes between lopsided approaches versus balanced ones. As organizations learn which behaviors drive the outcomes they wish to see within their residence halls, they can incorporate leadership training and development strategies to facilitate these outcomes.

Paladino, Murray, Newgent, and Gohn (2005) and Komives (1991) observed that supervisors' leadership behaviors contributed to environmental factors that could influenced RA satisfaction as well as burnout. When supervisors are effective leaders, they help those around them operate more effectively and feel better. When leaders are not effective, their limitations detrimentally impact the work environment. Thus,

counselors who help others learn to lead more effectively beneficially impact the individual as well as the surrounding environment.

Suggestions for Future Research

There are many potential avenues for future research. First, the LVI will benefit from additional refinement of sub-dimensions. The internal consistency can be improved and a bi-factor analysis will help clarify the factorial structure for the Forceful dimension. Since the primary function of this study was to develop and validate the LVI-S, there were few outcome variables included in this research design. Future research designs could pair the LVI-S with instruments such as the ACUHO-I/EBI Resident Assessment to investigate the relationship between leader versatility and residence hall outcomes. Rater source is another element that warrants further investigation. This study explored how RAs perceived the leadership characteristics of their direct supervisors, but multi-rater designs can incorporate more perspectives. Future research should include more rater sources, such as peers, the Hall Director's supervisor, and other personnel who may provide additional perspectives that could be valuable. The leadership behaviors a supervisor might value in a HD could markedly differ from the ones provided by an RA. For a true 360-degree perspective, additional rater sources should be incorporated into the study. Such designs would be even more robust if external effectiveness measures could be incorporated, such as the aforementioned climate surveys or internal performance evaluations.

Although the LVI-S was validated within a housing and residence-life context, a range of student affairs personnel and Higher Education faculty were involved the

development of the instrument. The item language is not residence-hall specific, so it could be used as a developmental feedback tool or research tool in other areas of student affairs. The concepts of Forceful and Enabling leadership appear to be applicable at many levels of leadership. They have been validated in an executive population as well as in an entry-level leadership population. Future studies could investigate how the instrument functions at higher levels of student affairs organizations, such as with assistant directors, directors, and higher-level administrators. Like the synergistic supervision model that preceded it, the LVI-S was conceived in a residence-hall context but its applicability could extend across the broader domain of student affairs.

Conversely, when the LVI-S was designed, the language was intended to be suitable for college student populations. Since students were involved during the validation process of the instrument, using the LVI-S with student populations is another avenue for potential research. Very few multi-rater leadership instruments have been validated within a student population. In addition, very few student leadership studies are conducted with multi-rater instruments. Most studies are conducted with self-rated leadership surveys, though a preponderance of evidence cautions against these methods. Self-rated leadership data is less reliable and exhibits limited predictive validity. The LVI-S can complement existing student leadership processes, helping to diversify the research methods used to measure college student leadership.

Conclusion

This study was designed to accomplish three objectives: 1) create a new leadership instrument that was designed and validated within a campus context; 2)

validate the concept of leadership versatility in a housing and residence life setting; and 3) to demonstrate that self-ratings of leadership performance have limited validity. Although the LVI-S will benefit from further refinement, results of the study indicated a promising start. Dimension-level internal consistency exceeded the threshold for fully developed psychometric instruments. Sub-dimension alphas were adequate for a four-item scale; but increasing sub-dimension alpha coefficients into the .70 and .80 range would enhance the performance of the instrument. Greater internal consistency will improve inter-rater reliability and will permit researchers to have greater confidence when interpreting sub-dimension analyses.

Comparisons with the SLPI, an established leadership instrument, provided evidence of convergent and discriminant validity. The Enabling dimension yielded 11 significant correlations with SLPI scales. The *Listens* and *Supports* sub-dimensions correlated significantly with all five SLPI scales ($p < .01$), and the *Empowers Others* sub-dimension correlated significantly with one scale. These correlations provided evidence of convergent validity in the Enabling dimension. The Forceful dimension exhibited seven fewer significant correlations with SLPI scales. In general, the magnitude of these correlations was lower and half of them were negative, providing evidence of discriminant validity. Both the Enabling and Forceful results were consistent with principles underlying the versatile leadership model.

Predictive validity of the LVI-S was strong, indicating versatility explained a considerable amount of the variance in effective residence hall leadership. The versatile leadership model proposed that leaders should be able to adjust their approach based on

the needs of the situation and the people involved. Versatile leaders were forceful, using authority to provide direction and clarify expectations; but this forcefulness was balanced by a supportive, considerate approach that enabled others to step up and deliver results. Thus, leadership versatility was an empirically valid approach for measuring leadership in a residence-life setting. Furthermore, the instrument aligns with a previously established model of residence hall supervision, the synergistic supervision model. Synergistic supervisors need to maintain focus on institutional responsibilities and uphold departmental directives, but they also need to build supportive, developmental relationships with supervisees. Effective supervisors find “synergy” between these two complementary, but opposing demands. The LVI-S measured the underlying behaviors necessary for synergistic supervision as well as how effectively leaders balanced between complementary opposites. No quantitative instrument has been able to measure behavioral synergy in a campus setting; therefore the predictive characteristics of the LVI-S may be interpreted as quantitative support for the synergistic supervision model as well as for leadership versatility. Qualitative research has demonstrated the value and impact of synergistic supervision, but the LVI-S could be used to extend this research in a quantitative direction.

The overall purpose of this study was to diversify the range of campus-based multi-rater instruments. An underlying premise was that self-ratings of leadership performance, absent other perspectives, were of questionable validity. Results from this study demonstrated that self-ratings were of dubious validity. First, self-ratings were unable to predict effectiveness. Second, with the exception of one sub-dimension, self-

ratings were unable to predict observers' ratings. Overall, the validity of self-ratings of performance should be questioned. This result underscored the need for a wider range of multi-rater instrumentation designed for and validated within a campus context.

REFERENCES

- ACUHO-I (2008) National Housing Training Institute request for proposal presentation
Retrieved February 27, 2009 from www.acuho.org/Portals/0/pdf/nhti_rfp_presentation.pdf
- Ames, D. R., & Flynn, F. J. (2007). What breaks a leader: The curvilinear relationship between assertiveness and leadership. *Journal of Personality and Social Psychology, 92*(2), 307-324
- Anotakis, J., Avolio, B. J., & Sivasubramaniam, N., (2003). Context and leadership: an examination of the nine-factor full-range leadership theory using the Multifactor Leadership Questionnaire. *The Leadership Quarterly, 14*, 261-295.
- Antonakis, J., Schriesheim, C. A., Donovan, J. A., Gopalkrishna-Pillai, K., Pellegrini, E. K., & Rossome, J. L. (2004). Methods for studying leadership. In J. A. Antonakis, A. T. Cianciolo, & R. J. Sternberg (Eds.), *The Nature of Leadership* (pp. 48-70). Thousand Oaks, CA: Sage Publications.
- Ansbacher, H. L., & Ansbacher, R. R., (Eds.) (1956). *The individual psychology of Alfred Adler*. New York: Basic Books.
- Aristotle (2004). *The Nicomachean ethics* (J. A. K. Thompson, Trans.). London: Penguin. (Original work published 350 B. C.)
- Arminio, J., & Creamer, D. G., (2001). What supervisors say about quality supervision. *College Student Affairs Journal, 21*(1), 35-44.

- Ashford S. (1989). Self-assessments in organizations: A literature review and integrative model. *Research in Organizational Behavior*, 11,33-174.
- Astin, A.W. (1993). *What matters in college: Four critical years revisited*. San Francisco: Jossey Bass.
- Astin, A. W. & Astin, H. S. (2000). *Leadership reconsidered: Engaging higher education in social change*. Battle Creek, MI: W. K. Kellogg Foundation.
- Astin, H. S. and Cress, C. (1998). *The impact of leadership programs on student development*. UCLA-HERI Technical Report to the W.K. Kellogg Foundation. Battle Creek, MI: W.K. Kellogg Foundation.
- Astin, A. W., & Sax, L. J. (1998). How undergraduates are affected by service participation. *Journal of College Student Development*, 39, 251-263.
- Atwater, L. E., & Yammarino, F. J. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, 45, 141-164.
- Avolio, B. J., Bass, B. M., & Jung, D. I. (1999). Re-examining the components of transformational and transactional leadership using the multifactor leadership questionnaire. *Journal of Occupational and Organizational Psychology*, 72, 441-462.
- Bales, R. F. (1950). *Interaction process analysis*. Reading, MA: Addison-Wesley.
- Bales, R. F. (1954). In conference. *Harvard Business Review*, 32(2), 44-50.
- Bass, B. M. (1985a). *Leadership and performance beyond expectations*. New York: The Free Press.

- Bass, B. M. (1985b). *Multifactor leadership questionnaire*. Binghamton, NY: State University of New York, School of Management.
- Bass, B. M. (1990). *Bass & Stogdill's Handbook of Leadership: Theory, Research, & Managerial Applications*. New York: The Free Press.
- Bass, B. M., & Avolio, B. J. (1990). *The multifactor leadership questionnaire*. Palo Alto, CA: Consulting Psychologists Press.
- Bass, B. M., & Avolio, B. J., (1994). *Improving organizational effectiveness through transformational leadership*. Thousand Oaks, CA: Sage.
- Bass, B. M., & Avolio, B. J., (1995/2000). Multifactor Leadership Questionnaire (2nd ed.) Technical Report. Redwood City, CA: Mindgarden, Inc.
- Bass, B. M., Avolio, B. J., Jung, D. I., & Berson, Y. (2003). Predicting unit performance by assessing transformational and transactional leadership. *Journal of Applied Psychology, 88*(2), 207-218.
- Bass, B. M., & Valenzi, E. R., (1974). Contingent aspects of effective management styles. In J. G. Hunt & L. L. Larsen (Eds.), *Contingency approaches to leadership*. Carbondale, IL: Southern Illinois University Press.
- Baxter, W. (2001). *Engineering leadership*. Unpublished master's thesis, Texas A&M University.
- Beehr, T. A., Ivanitskaya, L., Hansen, C. P., Erofeev, D., & Gudanowski, D. M. (2001). Evaluation of 360 degree feedback ratings: relationships with each other and with performance and selection predictors. *Journal of Organizational Behavior, 22*(7), 775-788.

- Billington, R. (1997). *Understanding Eastern Philosophy*. New York: Routledge.
- Birnbaum, P. H., Farh, J. L., & Wong, G. Y. (1986). The job characteristics model in Hong Kong. *Journal of Applied Psychology, 71*, 598-605.
- Blake, R. R., & McCauley, A. A. (1991). *Leadership dilemmas—Grid solutions*. Houston, TX: Gulf.
- Blake, R. R., & Mouton, J. S. (1964). *The managerial grid*. Houston, TX: Gulf.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for aggregation and analysis. In S. W. Kozlowski & K. J. Klein (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349-381). San Francisco: Jossey-Bass.
- Blimling, G. (1998). *The resident assistant: Applications and strategies for working with college students in residence halls* (5th ed.). Dubuque, Iowa: Kendall/Hunt Publishing Company.
- Blumberg, H. H. (1972). Communication of interpersonal evaluations. *Journal of Personality and Social Psychology, 23*, 157-162.
- Bono, J. E., & Judge, T. A. (2004). Personality and transformational and transactional leadership: A meta-analysis. *Journal of Applied Psychology, 89*(5), 901-910.
- Briscoe, J.P. and Hall, D.T. (1999) Grooming and picking leaders using competency frameworks: do they work? An alternative approach and new guidelines for practice. *Organizational Dynamics, 28*, 37–52.
- Brodsky, B. (1988). *Development of a modified version of the Leadership Practices Inventory for use with college students*. Unpublished master's thesis, San Jose

State University, California.

Brown, R. D. (1988). *Performance appraisal as a tool for staff development*. New

Directions for Student Services No. 43. San Francisco: Jossey-Bass.

Bureau of Labor Statistics. (2009, February 6). *Employment situation summary*. Retrived

February 8, 2009 from Bureau of Labor Statistics:

<http://www.bls.gov/news.release/empsit.nr0.htm>

Burns, J. M. (1978). *Leadership*. New York: Harper & Row.

Bycio, P., Hackett, R. D., & Allen, J. S., (1995). Further assessment of the Bass (1985)

conceptualization of transactional and transformational leadership. *Journal of Applied Psychology, 80*, 468-478.

Charan, R., Drotter, S., & Noel, J. (2001). *The leadership pipeline*. San Francisco:

Jossey-Bass.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition).

Hillsdale, NJ: Lawrence Earlbaum Associates.

Congressional Research Service. (2008, November 10). *The U. S. financial crisis: The global dimension with implications for U. S. Policy*. (RL34742) Washington, D.

C.: Congressional Research Service.

Conway, J. M., & Huffcutt, A. I. (1997). Psychometric properties of multisource

performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance, 10*(4), 331-360.

Creamer, D. G., & Janosik, S. M. (2003). Performance accountability that leads to

professional development. In S. M. Janosik, D. G. Creamer, J. B. Hirt, R. B.

- Winston, S. A. Saunders, & D. L. Cooper (Eds.) Supervising new professionals in student affairs: A guide for practitioners (pp. 43-84). New York: Brunner Routledge.
- Cress, C., Astin, H., Zimmerman-Oster, K., & Burkhardt, J. (2001). Developmental outcomes of college students' involvement in leadership activities. *Journal of College Student Development, 42*(1), 15-26.
- Crowne, D., & Marlowe, D. (1964). *The approval motive: Studies in evaluative independence*. New York: Wiley.
- Dalton, J. C. (1988) Employment and supervision of student affairs personnel. In M. J. Barr & Associates, *Student services and the law: A handbook for practitioners*. San Francisco: Jossey-Bass.
- Dalton, J. C. (1996) Managing human resources. In S. R. Komives, D. B. Woodard, Jr. & Associates, *Student Services: A handbook for the profession (3rd ed.)*. San Francisco: Jossey-Bass.
- Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35*, 867-881.
- DeLuga, R. J. (1989). The relationship between resident assistants' characteristics and leadership style and students' influencing behavior. *Journal of College and University Student Housing, 18*, 7-11.
- DeLuga, R. J. & Winters, J. J. (1991). Why the aggravation? Reasons students become resident assistants, interpersonal stress, and job satisfaction. *Journal of College Student Development, 32*, 546-552.

- Denison, D. R., Hoojberg, R., & Quinn, R. E. (1995). Paradox and performance: Toward a theory of behavioral complexity in managerial leadership. *Organization Science*, 6(5), 524-540.
- De Vader, C. L., Bateson, A. G., & Lord, R. G. (1986). Attribution theory: A meta-analysis of attributional hypotheses. In Locke E. (Ed.), *Generalizing from laboratory to field settings (pp. 63-81)* (pp. 63-81). Lexington, MA: Lexington Books.
- Dionne, S. D., Yammarino, F. J., Atwater, L. E., & James, L. R. (2002). Neutralizing substitutes for leadership theory: Leadership effects and common source bias. *Journal of Applied Psychology*, 87, 454-464.
- Dirkx, J. M., Gilley, J. W., Gilley, A. M. (2004). Change theory in CPE and HRD: Toward a holistic view of learning and change in work. *Advances in Developing Human Resources*, 6, 35-51.
- Dreikurs, R. (1971). *Social equality: The challenge of today*. Chicago, IL: Contemporary Books.
- Drucker, P. F. (1998). *On the profession of management*. Boston: Harvard Business School Publishing.
- Dugan, J. P. (2006a). Explorations using the social change model: Leadership development among college men and women. *Journal of College Student Development*, 47, 217-225.
- Dugan, J. P. (2006b). Involvement and leadership: A descriptive analysis of socially responsible leadership. *Journal of College Student Development*, 47(3), 335-343.

- Dugan, J. P., Garland, J. L., Jacoby, B., & Gasiorsky, A. (2008). Understanding commuter student self-efficacy for leadership: A within-group analysis. *NASPA Journal, 45*(2), 282-310.
- Dugan, J. P., & Komives, S. R. (2007). *Developing leadership capacity in college students: Findings from a national study*. A Report from the Multi-Institutional Study of Leadership. College Park, MD: National Clearinghouse for Leadership Programs.
- Easterling, D. & Foy, C. G. (2006). Trends in Greensboro's social capital: Preliminary findings from the 2006 National Social Capital Survey.
- Edwards, A. L. (1970). *The measurement of personality traits by scales and inventories*. New York: Holt, Rinehart, & Winston.
- Eich, D. J. (2008). *A grounded theory of high quality leadership programs: Perspectives from student leadership development programs in higher education*. Unpublished doctoral dissertation, University of Wisconsin, Madison.
- Eichinger, R. W. & Lombardo, M. M. (2003). Knowledge summary series: 360-degree assessment. *Human Resources Planning, 26* (4), 34-44.
- Ellevin, R. K., Allen, J., & Wircenski, M. (2001). Resident assistant training: A Southwestern perspective. *College Student Journal, 35*, 609-616.
- Fernandez, C. F., & Vecchio, R. P. (1997). Situational leadership theory revisited: A test of an across-jobs perspective. *Leadership Quarterly, 8*(1), 67-84.
- Fey C. (1991). *Mid-level student affairs administrators: A study of management skills and professional development needs*, Unpublished doctoral dissertation, Texas

A&M University, College Station.

- Fiedler, F. E. (1967). *A theory of leadership effectiveness*. New York: McGraw-Hill.
- Fiedler, F. E. (1993). The leadership situation and the black box in contingency theories. In M. M. Chemers & R. Ayman (Eds.), *Leadership theory and research: Perspectives and directions*. San Diego, CA: Academic Press.
- Fiedler, F. E. (1995). Reflections by an accidental theorist. *Leadership Quarterly*, 6(4), 453-461.
- French, J. R. P., Jr. & Raven, B. (1960). The bases of social power. In D. Cartwright & A. Zander (Eds.) *Group dynamics* (pp. 259-269). New York: Harper & Row.
- Friedman, T. L. (2007). *The world is flat: A brief history of the twenty-first century 3.0* (3rd ed.). New York: Picador.
- Gibb, C. A. (1947). The principles and traits of leadership. *Journal of Abnormal and Social Psychology*, 42, 267-284.
- Goffman, E. (1955). On face-work: An analysis of ritual elements in social interaction. *Psychiatry: Journal for the Study of Interpersonal Processes*, 18, 213-231.
- Gordon, S. E., Strode, C. B., & Mann, B.A. (1993). The mid-manager in student affairs: What are CSAOS looking for? *NASPA Journal*, 30(4), 290-297.
- Graeff, C. L. (1983). The situational leadership theory: A critical view. *Academy of Management Review*, 8, 285-291.
- Graeff, C. L. (1997). Evolution of situational leadership theory: A critical review. *Leadership Quarterly*, 8, 285-291.

- Graen, G. B., & Uhl-Bien, M. (1995). Relationship-based approach to leadership: Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level, multi-domain perspective. *Leadership Quarterly*, 6(2), 219-247..
- Greene, C. N., & Organ, D. W. (1973). An evaluation of causal models linking the received role with job satisfaction. *Administrative Science Quarterly*, 18, 95-103.
- Grote, D. (1996). *The complete guide to performance appraisal*. New York: American Management Association.
- Hall, D. T., & Moss, J. E. (1998). The new protean career contract: Helping organizations and employees adapt. *Organization Dynamics*, 26(3), 22-37.
- Harris, M. M., & Schaubroeck, J. (1988). A Meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43-62.
- Hersey, P., & Blanchard, K. H. (1969). Life cycle theory of leadership. *Training & Development Journal*, 23, 26-34.
- Hersey, P., & Blanchard, K. H. (1977). *Management of organizational behavior: Utilizing human resources (3rd ed.)*. Englewood Cliffs, NJ: Prentice-Hall.
- Hilzenrath, D. S., ElBoghdady, D. (2007, September 7). *Quarterly foreclosure rate again sets record*. Washington Post. Retrieved February 8, 2009, from <http://www.washingtonpost.com/wp-dyn/content/article/2007/09/06/AR2007090601137.html>
- Hogan, R., & Kaiser, R. B., (2005) What We Know About Leadership. *Review of General Psychology* 9(2), 169-180.

- Hollenbeck, G. P., & McCall, M. W. (2006). Leadership competency models [Theoretical and practitioner letters, letter 1]. *Leadership Quarterly, 17*, 398-413.
- Holmes, T. A. (1998). Performance-based approaches to human resource development. *New Directions for Student Services, 84*, 15-27.
- House, R. J. (1971). A path-goal theory of leader effectiveness. *Administrative Science Quarterly, 16*, 321-338.
- House, R. J. (1977). A theory of charismatic leadership. In J. G. Hunt & L. L. Larsen (Eds.), *Leadership: The cutting edge* (pp. 189-207). Carbondale, IL: Southern Illinois Press.
- House, R. J., & Aditya, R. N. (1997). The social scientific study of leadership: *Quo vadis?* *Journal of Management, 23*, pp 409-473.
- House, R. J., & Mitchell, T. R. (1974). Path-goal theory of leadership. *Journal of Contemporary Business, 3*, 81-97.
- Jaeger, A. J., & Caison, A. L. (2006). Rethinking criteria for training and selection: An inquiry into the emotional intelligence of resident assistants. *NASPA Journal, 43*(1), Art. 8. Retrieved February 15, 2009, from <http://publications.naspa.org/naspajournal/vol43/iss1/art8>
- Janosik, S. M., Creamer, D. G., Hirt, J. B., Winston, R. B., Saunders, S. A., & Cooper, D. L. (2003). *Supervising new professionals in student affairs: A guide for practitioners*. New York: Brunner-Routledge.
- Jenkins, W. O. (1947). A review of leadership studies with particular reference to military problems. *Psychological Bulletin, 44*, 54-79.

- Kaiser, R. B. (2005). *Filling the Leadership Pipeline*. Greensboro, N. C.: The Center for Creative Leadership Press.
- Kaiser, R. B., & Craig, S. B. (2001). Leadership Versatility Index technical report: Item selection and validation. Greensboro, NC: Kaplan DeVries Inc.
- Kaiser, R. B., Craig, S. B., Overfield, D., & Yarborough, J. P. (2009). Testing the leadership pipeline: Do the behaviors related to managerial effectiveness change with organizational level? Manuscript submitted for publication.
- Kaiser, R. B., Lindberg, J. T., Craig, S. B. (2007). Assessing the flexibility of managers: A comparison of methods. *International Journal of Selection and Measurement*, 15(1), 40-55.
- Kaplan, R.E. (1996). *Forceful leadership and enabling leadership: You can do both*. Greensboro, NC: Center for Creative Leadership.
- Kaplan, R. E., & Kaiser, R. B., (2003a). Developing versatile leadership. *MIT Sloan Management Review*, 44, 19-26.
- Kaplan, R.E. & Kaiser, R.B. (2003b). Rethinking a classic distinction in leadership: implications for the assessment and development of executives. *Consulting Psychology Journal: Research and Practice*, 55, 15-25.
- Kaplan, R. E., & Kaiser, R. B. (2006). *The versatile leader: Make the most of your strengths—without overdoing it*. San Francisco: Pfeiffer.
- Kahn, R. L., & Katz, D. (1953). Leadership practices in relation to productivity and morale. In D. Cartwright & A. Zander (Eds.), *Group Dynamics*. New York: Harper & Rowe.

- Kane, N. E. (1982). A comparative study of student affairs mid-level professionals: Characteristics, perceived skill attainment and need for continued development (Doctoral dissertation, Florida State University, 1982). *Dissertation Abstracts International*, 43(9), 2902A.
- Kluger, A. N. and DeNisi, A. (1996) The effects of feedback interventions on performance: historical review, a meta-analysis and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284.
- Kodres, L. (2008). A crisis of confidence ... and a lot more [electronic version]. *Finance & Development* 45(2). Retrieved February 8, 2009, from International Monetary Fund: <http://www.imf.org/external/pubs/ft/fandd/2008/06/kodres.htm>
- Komives, S. R. (1991a). Gender differences in the relationship of hall directors' transformational and transactional leadership and achieving styles. *Journal of College Student Development*, 32, 155-164.
- Komives, S. R. (1991b). The relationship of hall directors' transformational and transactional leadership to select resident assistant outcomes. *Journal of College Student Development*, 32, 509-515.
- Komives, S. R. (1991c). The relationship of same- and cross-gender work pairs to staff performance and supervisor leadership in residence hall units. *Sex Roles*, 24, 355-363.
- Komives, S. R., Longerbeam, S. D., Owen, J. O., Mainella, F., & Osteen, L. (2006). A leadership identity development model: Applications from a grounded theory. *Journal of College Student Development*, 47(4), 401-418.

- Komives, S. R., Lucas, N., & McMahon, T. R. (2006). *Exploring leadership: For college students who want to make a difference (2nd edition)*. San Francisco: Jossey-Bass.
- Komives, S. R., Owen, J. O., Longerbeam, S. D., Mainella, F., & Osteen, L. (2005). Developing a leadership identity: A grounded theory. *Journal of College Student Development, 46*(6), 593-611.
- Kotter, J. P. (1990). *A force for change: How leadership differs from management*. New York: The Free Press.
- Kouzes, J., & Posner, B. Z. (1987). *The leadership challenge: How to get extraordinary things done in organizations*. San Francisco: Jossey-Bass.
- Kouzes, J., & Posner, B. Z. (2002). *The leadership challenge (3rd ed.)*. San Francisco: Jossey-Bass.
- Kouzes, J. & Posner, B. Z. (2003). *The Student Leadership Practices Inventory*. San Francisco: Jossey-Bass.
- Kraut, R. (2008) *Aristotle's Ethics*, The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), E N. Zalta (Ed.), Retrieved February 8, 2009, from <http://plato.stanford.edu/archives/fall2008/entries/aristotle-ethics/>
- Kroeck, K. G., Lowe, K. B., & Brown, K. W. (2004). The assessment of leadership. In J. A. Antonakis, A. T. Cianciolo, & R. J. Sternberg (Eds.), *The Nature of Leadership* (pp. 48-70). Thousand Oaks, CA: Sage Publications.
- Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology, 67*(4), 596-610.

- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134.
- Likert, R. (1961). *New Patterns of Management*. New York: McGraw-Hill.
- LeBreton, J. M., Burgess, J. R. D., Kaiser, R. B., Atchley, E. K., & James, L. R. (2003). The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar? *Organizational Research Methods*, 6, 78-126.
- Lewin, K., & Lippett, R. (1938). An experimental approach to the study of autocracy and democracy: A preliminary note. *Sociometry*, 1, 292-300.
- Leslie, J. B., & Fleenor, J. W. (1998). *Feedback to managers: A review and comparison of multi-rater instruments for management development*. Greensboro, NC: Center for Creative Leadership.
- Likert, R. (1932). "A technique for the measurement of attitude scales." *Archives of Psychology*, 140, 44-53.
- Likert, R. (1967). *The human organization*. New York: McGraw-Hill.
- Lipman-Blumen, J. (1987). *Achieving styles: A handbook for researchers and human resource professionals*. Claremont, CA: L-BLA Associates.
- Lipman-Blumen, J., & Leavitt, H. J. (1979). *L-BLA achieving styles inventory, form 13*. Bethesda, MD: L-BLA Associates.
- Lombardo, M. M., & Eichinger, R. W. (2000). *The leadership machine*. (2nd ed.). Minneapolis, MN: Lominger Limited, Inc.

- Lord, R. G., Foti, R. J., & De Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance*, 34, 343-378.
- Lowe, K. B., Kroeck, K. G., & Sivasubraminiam, N., (1996). Effectiveness correlates of transformation and transactional leadership: A meta-analytic review of the MLQ literature. *Leadership Quarterly*, 7(3), 385-425.
- Lowe, K. B., & Gardner, W. L. (2000). Ten years of *The Leadership Quarterly*: Contributions and challenges for the future. *Leadership Quarterly*, 11, 459-514.
- Lovell, C. D., & Kosten, L. A. (2000). Skills, knowledge and personal traits necessary for success as a student affairs administrator: A meta-analysis of thirty years of research. *NASPA Journal*, 37(4), 553-572.
- Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), 280-296.
- Machlup, F. 1962. *The production and distribution of knowledge in the United States*. Princeton: Princeton University Press.
- Mann, F. C. (1965). Toward an understanding of the leadership role in formal organization. In R. Dublin (Ed.), *Leadership and Productivity* (pp. 68-103). San Francisco: Chandler.
- Mann, R. D. (1959). A review of the relationship between personality and performance in small groups. *Psychological Bulletin*, 56, 241-270.
- Martin, R. (2007). How successful leaders think. *Harvard Business Review*, 85, 60-67.
- Martinko, M. J., Harvey, P., & Douglas, S. C., (2007). The role, function, and

- contribution of attribution theory to leadership: A review. *Leadership Quarterly*, 18, 561-585.
- Matlin, M., & Stang, D. (1978). *The Pollyanna principle*. Cambridge, MA: Schenkman.
- McCall, M. W., Jr., & Lombardo, M. M. (1983). *Off the track: Why and how successful executives get derailed*. Greensboro, NC: Center for Creative Leadership.
- McCall, M. W., Jr., & Segrist, S. A., (1978). *In pursuit of the manager's job: Building on Mintzberg* (Tech. Rep. No. 14). Greensboro, NC: Center for Creative Leadership.
- McCauley, C.D. & Van Velsor, E. (2004). *The center for creative leadership handbook of leadership development. (2nd ed.)*. San Francisco, CA: Jossey-Bass.
- Mintzberg, H. (1973). *The nature of managerial work*. New York: Harper & Row.
- Misumi, J. (1985). *The behavioral science of leadership: An interdisciplinary Japanese research program*. Ann Arbor: University of Michigan Press.
- Morphew, C. C., & Hartley, M. (2006, January). Mission Statements: A Thematic Analysis of Rhetoric Across Institutional Type. *Journal of Higher Education*, 77, 456-471.
- Murray, J. L., Snider, B. R., & Midkiff, R. M. (1999). The effects of training on resident assistant job performance. *Journal of College Student Development*, 40 (6), 744-747.
- The National Conference on Citizenship. (2006, September 18). *America's civic health index: Broken engagement*. Washington, D. C.: Author.
- Northouse, P. G. (2004). *Leadership Theory and Practice* (3rd ed.). Thousand Oaks, CA: Sage.

- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Organ, D. W., & Greene, C. N. (1981). The effects of formalization on professional involvement: A compensatory process approach. *Administrative Science Quarterly*, 26, 237-252.
- Paladino, D. A., Murray, T. L., Newgent, R. A., Gohn, L. A. (2005). Resident assistant burnout: Factors impacting depersonalization, emotional exhaustion, and personal accomplishment. *Journal of College and University Housing*, 33, 19-27.
- Parry, K. W. & Sinha, P. N. (2005). Researching the trainability of transformational organizational leadership. *Human Resource Development International*, 8, 165-183.
- Pascarella, E.T., & Terenzini, P.T. (2005). *How college affects students. Findings and insights from twenty years of research*. San Francisco: Jossey-Bass.
- Peters, L. H., Hartke, D. D., & Pohlman, J. T. (1985). Fiedler's contingency theory of leadership: An application of the meta-analysis procedures of Schmidt and Hunter. *Psychological Bulletin*, 97, 274-285.
- Pike, G. R. (2000). The influence of fraternity or sorority membership on students' college experiences and cognitive development. *Research in Higher Education*, 41, 117-139.
- Podsakoff, P. M., MacKenzie, S. B., Moorman, R. H., & Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *Leadership Quarterly*, 1, 107-142.

- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12(4), 531-544.
- Posner, B. Z. (2004). A leadership development instrument for students: Updated. *Journal of College Student Development*, 45, 443-456.
- Posner, B. Z., & Brodsky, B. (1993). The leadership practices of effective RAs. *Journal of College Student Development*, 34(4), 300-304.
- Posner, B., & Rosenberger, J. (1997). Effective orientation advisors are leaders too. *NASPA Journal*, 35(1), 46-56.
- Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.
- Quinn, R. E. (1988) *Beyond rational management*. San Francisco: Jossey-Bass.
- Roe, B. B. (1981, March). *An analysis of career enrichment needs and programs for student affairs personnel*. Paper presented at the Annual Convention of the American Personnel and Guidance Association, Detroit, MI.
- Roberts, D. M. (2007). Preferred methods of professional development in student affairs. *NASPA Journal*, 44, 561-577.
- Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attributional processes. *Journal of Experimental Social Psychology*, 13, 279-301.
- Rost, J. C. (1993). *Leadership for the 21st century*. Westport, CT: Praeger.
- Salancik, G. R., & Pfeffer, J. (1977). An examination of need-satisfaction models of job attitudes. *Administrative Science Quarterly*, 22, 427-456.

- Sashkin, M. (1984). *The visionary leader: The Leader Behavior Questionnaire*. Bryn Mawr, PA: Organization Design and Development.
- Sashkin, M. (2004). *Transformational leadership approaches: A review and synthesis*. In J. A. Antonakis, A. T. Cianciolo, & R. J. Sternberg (Eds.), *The Nature of Leadership* (pp. 48-70). Thousand Oaks, CA: Sage Publications.
- Schon, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Schreiber, P. J., Dunkel, N. W., & Jahr, P. K. (1994). Competency based developmental programs. *The College Student Affairs Journal*, 14(1), 23–30.
- Schriesheim, C. A., Castro, S. L., Zhou, X., & Yammarino, F. J., (2001) The folly of theorizing “A” but testing “B” A selective level-of-analysis review and a detailed Leader-Member Exchange illustration. *Leadership Quarterly*, 12, 515-551.
- Schriesheim, J. R., & Schriesheim, C. A. (1980). A test of the path-goal theory of leadership and some suggested directions for future research. *Personnel Psychology*, 33, 349-370.
- Schwarz, N. (1999). Self reports: How the questions shape the answers. *American Psychologist*, 54, 93-105.
- Schwartz, M., & Gimbel, K. (2000). *Leadership resources: A guide to training and development tools*. Greensboro, NC: Center for Creative Leadership.
- Sheridan, J., & Vredenburg, D. (1978). Predicting leader behavior in a hospital organization. *Academy of Management Journal*, 21,679-689.
- Shertzer, J. E., & Schue, J. H. (2004). College student perceptions of leadership:

- Empowering and constraining beliefs. *NASPA Journal*, 42, 111-131.
- Shipper, F. (1991). Mastery and frequency of managerial behaviors relative to sub-unit effectiveness. *Human Relations*, 44(4), 371-388.
- Shrout, P. & Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- Silzer, R. F. (2006). Leadership competency models [Theoretical and practitioner letters, letter 2]. *Leadership Quarterly*, 17, 398-413.
- Smart, J. C., Ethington, C. A., Riggs, R. O., & Thompson, M. D. (2002). Influences of institutional expenditure patterns on the development of students' leadership competencies. *Research in Higher Education*, 43(1), 115-132.
- Snyder, C. R., Higgins, R. L., & Stucky, R. J. (1983). *Excuses: Masquerades in search of grace*. New York: Wiley.
- Snyder, C. R., Shenkel, R. J., & Lowery, C. R. (1977). Acceptance of personality interpretations: The "Barnum effect" and beyond. *Journal of Consulting and Clinical Psychology*, 45, 104-114.
- Standards for educational and psychological testing. (1985). Washington, DC: American Psychological Association.
- Stockford, L. & Bissell, H.W. (1949). Factors involved in establishing a merit-rating scale. *Personnel*, 26, 94-116.
- Stogdill, R. M. (1948). Personal factors associated with leadership: A survey of the literature. *Journal of Psychology*, 25, 35-71.

- Stogdill, R. M. & Coons, A. E. (1957). *Leader Behavior: Its description and measurement*. Columbus, OH: Ohio State University Press for Business Research.
- Strube, M. J., & Garcia, J. E. (1981). A meta-analytic investigation of Fiedler's contingency model of leadership effectiveness. *Psychological Bulletin*, 90, 307-321.
- The Student Leadership Practices Inventory, (n. d.). Retrieved February 28, 2009 from <http://www.leadershipchallenge.com/WileyCDA/LCTitle/productCd-PCOL44.html>
- Swanbrow, D. (2007, August 7). *Wealth gap is increasing, University of Michigan study shows*. Retrieved February 8, 2009, from University of Michigan: <http://www.ns.umich.edu/htdocs/releases/story.php?id=5981>
- Teegarden, P. H. (2004). *Change ahead: The 2004 nonprofit executive leadership and transitions survey*. Baltimore: Annie E. Casey Foundation.
- Tejada, M. J., Scandura, T. A., Pillai, R. (2001). The MLQ revisited: Psychometric properties and recommendations. *Leadership Quarterly*, 12, 31-52.
- Tepper, B. J., Percy, P. M. (1994). Structural validity of the Multifactor Leadership Questionnaire. *Educational and Psychological Measurement*, 54, 734-744.
- Tesser, A., & Rosen, S. (1975). The reluctance to transmit bad news. In L. Berkowitz (Ed.), *Advances in experimental psychology* (Vol. 8, pp. 193-232). New York: Academic Press.

- Thompson, M. D. (2006). Student leadership process development: An assessment of contributing college resources. *Journal of College Student Development, 47*, 343-350.
- Tierney, T. J. (2006). *The nonprofit's sector's leadership deficit*. (white paper). San Francisco: The Bridgespan Group.
- Tillotson, E. A. (1995). *An analysis of technical, human and conceptual skills among student affairs administrators in higher education*. Unpublished doctoral dissertation, Texas A&M University, College Station.
- Turrentine, C. G. (2001). A comparison of self-assessment and peer assessment of leadership skills. *NASPA Journal, 36*(3) 361-371.
- Tsui, A. S. (1984). A role set analysis of managerial reputation. *Organizational Behavior and Human Performance, 34*, 64-96.
- Tyree, T. M. (1998). Designing an instrument to measure the socially responsible leadership using the social change model of leadership development. *Dissertation Abstracts International, 59*(06), 1945. (AAT 9836493).
- Upcraft, M. L. & Pilato, G. T. (1982). *Residence hall assistants in college: A guide to selection, training, and supervision*. San Francisco: Jossey Bass.
- U. S. Treasury Department. (2009). *The daily history of the debt results*. Retrieved February 8, 2009, from TreasuryDirect Web site:
<http://www.treasurydirect.gov/NP/NPGateway>.
- Vecchio, R. P. (1987). Situational leadership theory: An examination of a prescriptive theory. *Journal of Applied Psychology, 72*(3), 444-451.

- Vecchio, R. P. (2007). Situational leadership theory: An examination of a prescriptive view. In R. P. Vecchio (Ed.) *Understanding the dynamics of power and influence in organizations*. Notre Dame, IN: University of Notre Dame.
- Vecchio, R. P., & Boatwright, K. J. (2002). Preference for idealized style of supervision. *Leadership Quarterly*, 13, 327-342.
- Vroom, V. H., & Yetton, P. W. (1973). *Leadership and decision-making*. Pittsburgh, PA: University of Pittsburgh Press.
- Weik, P. M. (2005). Building the executive ranks: Current practices in developing future business leaders. In R. B. Kaiser (Ed.), *Filling the Leadership Pipeline* (pp. 7-23). Greensboro, NC: Center for Creative Leadership.
- Wicklund R. (1975). Objective self-awareness. In Berkowitz L (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 233-275). New York: Academic Press.
- Wicklund R. (1978). Three years later. In Berkowitz L (Ed.), *Cognitive theories in social psychology* (pp. 509-521). New York: Academic Press.
- Wicklund R. (1979). The influence of self on human behavior. *American Scientist*, 67, 187-193.
- Winston, R. B. & Creamer, D. G. (1997). *Improving staffing practices in student affairs*. San Francisco: Jossey-Bass.
- Winston, R. B. & Creamer, D. G. (1998). Staff supervision and professional development: An integrated approach. *New Directions for Student Services*, 84, 29-42.

- Winston, R. B. & Fitch, R. T. (1993). In R. Winston & Scott Anchors and Associates (Eds.), *Student housing and residential life: A handbook for professionals committed to student development goals* (pp. 315-343). San Francisco: Jossey Bass.
- Winston, R. B., & Hirt, J. B. (2003). Activating synergistic supervision approaches: Practical suggestions. In S. M. Janosik, D. G. Creamer, J. B. Hirt, R. B. Winston, S. A. Saunders, & D. L. Cooper (Eds.) *Supervising new professionals in student affairs: A guide for practitioners* (pp. 43-84). New York: Brunner Routledge.
- Winston, R. B., Ullom, M. S., Werring, C. J. (1984). Student paraprofessionals in residence halls. In S. C. Ender and R. B. Winston, Jr. (Eds.) *Students as paraprofessional staff*. New Directions for Student Services, no. 27. San Francisco: Jossey-Bass.
- Woffard, J. C., & Liska, L. Z. (1993). Path-goal theories of leadership: A meta-analysis. *Journal of Management*, 19(4), 857-876.
- Yammarino, F. J., Dionne, S. D., Chun, J. U., & Dansereau, F. (2005). Leadership and levels of analysis: A state of the science review. *Leadership Quarterly*, 16(6), 879-919.
- Yukl, G. (1989). *Leadership in organizations* (2th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Yukl, G. (1994). *Leadership in organizations* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.
- Yukl, G. (1998). *Leadership in organizations* (4th ed.). Upper Saddle River, NJ: Prentice

Hall.

Zaccaro, S. J., Foti, R. J. and Kenny, D. A. (1991) Self monitoring and trait-based variance in leadership: an investigation of leader flexibility across multiple group situations. *Journal of Applied Psychology*, 76, 308–315.

Zaccaro, S. J., Gilbert, J. A., Thor, K. K. and Mumford, M. D. (1991) Leadership and social intelligence: linking social perceptiveness and behavioral flexibility to leader effectiveness. *Leadership Quarterly*, 2, 317–342.

Zaleznik, A. (1974). Charismatic and consensual leaders: A psychological comparison. *Bulletin of the Menninger Clinic*, 38, 222-238.

Zimmerman-Oster, K. & Burkhardt, J. (1999a). *Leadership in the making: Impact and insights from leadership development programs in U.S. colleges and universities*. Battle Creek, MI: W.K. Kellogg Foundation.

Zimmerman-Oster, K., & Burkhardt, J.C. (1999b). Leadership in the making: A comprehensive examination of the impact of leadership development programs on students. *The Journal of Leadership Studies*, 6(3/4), 50-66.

LVI-S Hall Director / Supervisor Response Form

Demographic Questionnaire

****If you do not supervise RAs, please click the "back" button on your browser to return to the previous page, then select the Resident Advisor link.****

Hd1. **Enter your university or college-issued e-mail address. Omit the "@" symbol and anything that follows.** For example, jpyarbor@uncg.edu would be entered as jpyarbor). This information will be encrypted.

Hd2. Placeholder item to preserve formatting; no answer required

Hd3. Which institution are you affiliated with?

- University of North Carolina at Greensboro
- University of North Carolina at Chapel Hill
- University of North Carolina at Charlotte
- Appalachian State University
- Guilford College
- Greensboro College
- Wake Forest University
- North Carolina State University
- North Carolina A&T

Hd4. Gender

- Male
- Female

Hd5. Date of Birth month/day/year (XX/XX/XXXX)

Hd6. Racial affiliation

- African-American
- White
- Latino/a
- Native American
- Asian
- Pacific Islander

- Bi-/Multi-racial
- Other

Hd6. Number of years I've been a Hall Director (Resident Coordinator, etc.)

Hd7. Were you an RA as an undergraduate?

- Yes
- No

Hd8. If you answered Yes to the previous question, how many years were you an RA?

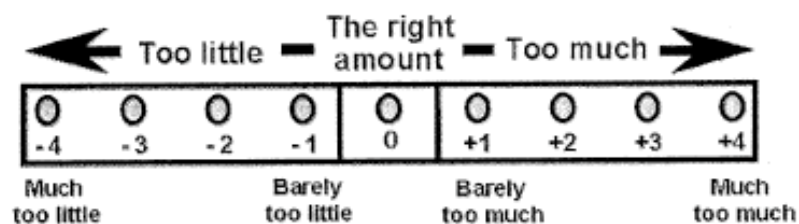
Hd9. The RAs I supervise work in a

- Traditional Residence Hall
- Learning Community
- Combination of both _____
- Other (if "other" please specify _____)

Hd10. How many RAs do you supervise?

Please read the following information very carefully.

The rating scale used in the next section is different from the typical kind, where a higher score is a better score. **On this scale, the best score is a "0," in the middle of the scale.** The premise is that there are two kinds of performance problems: when leaders emphasize something too much or when they put too little emphasis on something.



WARNING!: Some people misread this scale. Please do not mistake it for the usual type where a high score is the best score.

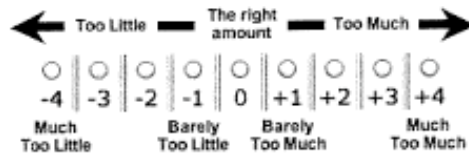
You might think about this rating scale as operating similar to the Goldilocks principle:

1. Use the "Too much" side of the scale for behaviors that you take to an extreme—what you do too frequently or with too much intensity.

with too little intensity.

3. Use the "The right amount" for items that you do to the correct degree—neither too much nor too little.

Rate your leadership characteristics.



| | | | | | | | | | |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. Take charge—in control of my area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. Take the initiative—eager to lead. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. Give clear direction—tell people what to do. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. Step in—get actively involved when problems arise. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Decisive—make up my mind quickly. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Let people know clearly where I stand on issues—declare myself. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. Tell people what is on my mind—forthcoming. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Defend my point-of-view—don't back down easily. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. Challenge people to do their best—push hard for high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. Expect a lot—insist on high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. Direct—tell people when I am dissatisfied with their work. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. Hold people accountable—am firm when others do not meet my standards. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | | | | | | | | | |
| 13. Empower other people—able to step back and give them room to do their job. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 14. Give people plenty of opportunity to show initiative. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

| | | | | | | | | | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 15. Give people freedom to decide how to perform their duties—hands-off. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 16. Trust people to handle the problems that come up in their area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 17. Participative—include people when making decisions. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 18. Want to know where others stand—ask for others' opinions. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 19. Can be persuaded to change my mind—open to influence. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 20. Make it safe to challenge or critique my thinking—welcome dialogue and debate. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 21. Show appreciation—go out of my way to make others feel good about their contribution. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 22. Friendly—am nice and treat other people with courtesy. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 23. Sensitive—careful not to hurt the other person's feelings. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 24. Cut people slack—am understanding when others are not able to do their duties and responsibilities. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| | | | | | | | | | |

The next set of questions uses a different scale, so please read carefully. How frequently do you typically engage in the following behaviors and actions? Select the button to the right of each statement, using the scale below, that best applies.

Rarely Or Seldom
 Once In A While
 Sometimes
 Very Often
 Frequently

| | Rarely or Seldom | Once In A While | Sometimes | Very Often | Frequently |
|---|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 25. Sets a personal example of what I expect from other people. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |

| | | | | | |
|---|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 26. Look ahead and communicate about what I believe will affect us in the future. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 27. Look around for ways to develop and challenge my skills and abilities. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 28. Foster cooperative rather than competitive relationships among people I work with. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 29. Praise people for a job well done. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 30. Spend time and energy making sure that people in our organization adhere to the principles and standards we have agreed on. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 31. Describe to others in our organizations what we should be capable of accomplishing. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 32. Look for ways that others can try out new ideas and methods. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 33. Actively listen to diverse points of view. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 34. Encourage others as they work on activities and programs in our organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 35. Follow through on the promises and commitments I makes in this organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 36. Talk with others about sharing a vision of how much better the organization could be in the future. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 37. Keep current on events and activities that might affect our organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 38. Treat others with dignity and respect. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 39. Give people in our organization support and expresse appreciation for their contribution. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 40. Find ways to get feedback about how my actions affect other people's performance. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 41. Talk with others about how their own interests can be met by working toward a common goal. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |

| | | | | | |
|--|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 42. When things do not go as we expected, I ask, "What can we learn from this experience?" | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 43. Support the decisions that other people in our organization make on their own. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 44. Make it a point to publicly recognize people who show commitment to our values. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 45. Build consensus on an agreed-on set of values for our organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 46. Am upbeat and positive when talking about what our organization aspires to accomplish. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 47. Make sure that we set goals and make specific plans for the projects we undertake. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 48. Give others a great deal of freedom and choice in deciding how to do their work. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 49. Find ways for us to celebrate accomplishments. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 50. Talk about the values and principles that guide my actions. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 51. Speak with conviction about the higher purpose and meaning of what we are doing. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 52. Take initiative in experimenting with the way we can do things in our organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 53. Provide opportunities for others to take on leadership responsibilities. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 54. Make sure that people in our organization are creatively recognized for their contributions. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |

Please click the "submit" button below to enter your responses.

Submit

The Leadership Versatility Index - student version is derived from the Leadership Versatility Index: Copyright© 2006 Kaplan DeVries Inc. All rights reserved. The method of assessment used by the Leadership Versatility Index® is protected by U.S. Patent No. 7,121,830. Student version adapted with permission from Kaplan DeVries Inc.

In reference to items 25-54: Copyright© 2003 James M. Kouzes and Barry Z. Posner. All

S4. Make sure that people in our organization are creatively recognized for their contributions.

1

2

3

4

5

Please click the "submit" button below to enter your responses.

Submit

The Leadership Versatility Index - student version is derived from the Leadership Versatility Index: Copyright© 2006 Kaplan DeVries Inc. All rights reserved. The method of assessment used by the Leadership Versatility Index® is protected by U.S. Patent No. 7,121,830. Student version adapted with permission from Kaplan DeVries Inc.

In reference to items 25-54: Copyright© 2003 James M. Kouzes and Barry Z. Posner. All

10/23/09

Document

<http://versatileleadershipstudy.com/HallDirectorFor>

rights reserved. Used with permission.

APPENDIX B. RESIDENT ADVISOR RESPONSE FORM USED FOR MAIN STUDY

LVI-S Resident Advisor Response Form

Demographic Questionnaire

****If you are not an RA, please click the "back" button on your browser to return to the previous page, then select the Hall Director/Supervisor link.****

Ra1. Enter YOUR SUPERVISOR'S university or college-issued e-mail address. Omit the "@" symbol and anything that follows. For example, *jpyarbor@uncg.edu* would be entered as *jpyarbor*). This information will be encrypted. --->

Ra2. Enter YOUR university or college-issued e-mail address. Omit the "@" symbol and anything that follows. For example, *jpyarbor@uncg.edu* would be entered as *jpyarbor*). This information will be encrypted. --->

Ra3. Which institution are you affiliated with?

- University of North Carolina at Greensboro
- University of North Carolina at Chapel Hill
- University of North Carolina at Charlotte
- Appalachian State University
- Guilford College
- Greensboro College
- Wake Forest University
- North Carolina State University
- North Carolina A&T

Ra4. Gender

- Male
- Female

Ra5. Date of Birth month/day/year (XX/XX/XXXX)

Ra6. Racial affiliation

- African-American
- White
- Latino/a
- Native American
- Asian
- Pacific Islander

- BI-/Multiracial
- Other

Ra7. Year in school

- Sophomore
- Junior
- Senior

Ra8. Number of years I've worked as an RA

- One year
- Two years
- Three years
- Four years

Ra9. My floor is considered a

- Traditional Residence Hall
- Learning Community
- Other (if "other" please specify: _____)

Ra10. Approximate number of residents living on your hall _____

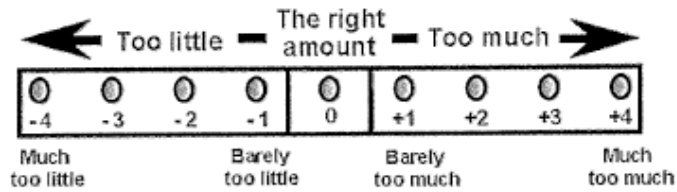
In the remainder of this survey, you will be asked to rate the leadership characteristics of your immediate supervisor. This is the person to whom you report and who has direct supervisory responsibility for you.

I have worked with my supervisor enough to form an opinion on his/her leadership style.

** (Please click the drop-down menu to select a response) ** --> ▾

Please read the following information very carefully.

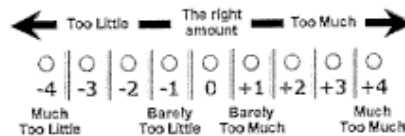
*The rating scale used in the next section is different from the typical kind, where a higher score is a better score. **On this scale, the best score is a "0," in the middle of the scale.** The premise is that there are two kinds of performance problems: when leaders emphasize something too much or when they put too little emphasis on something.*



WARNING! Some people misread this scale. Please do not mistake it for the usual type where a high score is the best score.

You might think about this rating scale as operating similar to the Goldilocks principle:

1. Use the "Too much" side of the scale for behaviors that your supervisor takes to an extreme—what he/she does too frequently or with too much intensity.
2. Use the "Too little" side for those behaviors that your supervisor is deficient on—what she/he does not do often enough or does with too little intensity.
3. Use the "The right amount" for items that the leader does to the correct degree, neither too much, nor too little.



| | | | | | | | | | |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. Takes charge—in control of her/his area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 2. Takes the initiative—eager to lead. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 3. Gives clear direction—tells people what to do. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 4. Steps in—gets actively involved when problems arise. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 5. Decisive—makes up his/her mind quickly. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 6. Lets people know clearly where he/she stands on issues—declares her-/himself. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 7. Tells people what is on her/his mind—forthcoming. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 8. Defends his/her point-of-view—doesn't back down easily. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 9. Challenges people to do their best—pushes hard for high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |

| | | | | | | | | | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 10. Expects a lot—Insists on high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 11. Direct—tells people when she/he is dissatisfied with their work. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 12. Holds people accountable—is firm when others do not meet his/her standards. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| <p>← Too Little The right amount Too Much →</p> <p><input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>-4 -3 -2 -1 0 +1 +2 +3 +4</p> <p>Much Barely Barely Much</p> <p>Too Little Too Little Too Much Too Much</p> | | | | | | | | | |
| 13. Empowers other people—able to step back and give them room to do their job. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 14. Gives people plenty of opportunity to show initiative. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 15. Gives people freedom to decide how to perform their duties—hands-off. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 16. Trusts people to handle the problems that come up in their area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 17. Participative—includes people when making decisions. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 18. Wants to know where others stand—asks for others' opinions. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 19. Can be persuaded to change her/his mind—open to influence. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 20. Makes it safe to challenge or critique his/her thinking—welcomes dialogue and debate. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 21. Shows appreciation—goes out of his/her way to make others feel good about their contribution. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 22. Friendly—is nice and treats other people with courtesy. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 23. Sensitive—careful not to hurt the other person's feelings. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 24. Cuts people slack—is understanding when they are not able to do their duties and responsibilities. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| <p>← Too Little The right amount Too Much →</p> <p><input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/></p> <p>-4 -3 -2 -1 0 +1 +2 +3 +4</p> <p>Much Barely Barely Much</p> <p>Too Little Too Little Too Much Too Much</p> | | | | | | | | | |

The next set of questions uses a different scale, so please read carefully. How frequently does this person typically engage in the following behaviors and actions? Select the button to the right of each statement, using the scale below, that best applies.

Rarely Or Seldom
 Once In A While
 Sometimes
 Very Often
 Frequently

| | Rarely or Seldom | Once In A While | Sometimes | Very Often | Frequently |
|--|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 25. Sets a personal example of what he or she expects from other people. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 26. Looks ahead and communicates about what he or she believes will affect us in the future. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 27. Looks around for ways to develop and challenge his or her skills and abilities. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 28. Fosters cooperative rather than competitive relationships among people he or she works with. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 29. Praises people for a job well done. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 30. Spends time and energy making sure that people in our organization adhere to the principles and standards we have agreed on. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 31. Describes to others in our organizations what we should be capable of accomplishing. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 32. Looks for ways that others can try out new ideas and methods. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 33. Actively listens to diverse points of view. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 34. Encourages others as they work on activities and programs in our organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 35. Follows through on the promises and commitments he or she makes in this organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 36. Talks with others about sharing a vision of how much better the organization could be in the future. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |

| | | | | | |
|---|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 37. Keeps current on events and activities that might affect our organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 38. Treats others with dignity and respect. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 39. Gives people in our organization support and expresses appreciation for their contribution. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 40. Finds ways to get feedback about how his or her actions affect other people's performance. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 41. Talks with others about how their own interests can be met by working toward a common goal. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 42. When things do not go as we expected, asks, "What can we learn from this experience?" | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 43. Supports the decisions that other people in our organization make on their own. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 44. Makes it a point to publicly recognize people who show commitment to our values. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 45. Builds consensus on an agreed-on set of values for our organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 46. Is upbeat and positive when talking about what our organization aspires to accomplish. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 47. Makes sure that we set goals and make specific plans for the projects we undertake. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 48. Gives others a great deal of freedom and choice in deciding how to do their work. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 49. Finds ways for us to celebrate accomplishments. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 50. Talks about the values and principles that guide his or her actions. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 51. Speaks with conviction about the higher purpose and meaning of what we are doing. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 52. Takes initiative in experimenting with the way we can do things in our organization. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |
| 53. Provides opportunities for others to take on leadership responsibilities. | <input type="radio"/> 1 | <input type="radio"/> 2 | <input type="radio"/> 3 | <input type="radio"/> 4 | <input type="radio"/> 5 |

54. Makes sure that people in our organization are creatively recognized for their contributions.

1

2

3

4

5

The final three questions address your opinion of your supervisor's overall effectiveness.

55. Overall, to what extent do you believe your supervisor is performing his/her job the way you would like it to be performed? Moderately (about 50% of the time)

56. To what extent has she/he met your own expectations in her/his administrative roles and responsibilities? Moderately (about 50% of the time)

57. If you had entirely your own way, to what extent would you change the manner in which he/she is doing the job? Moderately (about 50% of the time)

Please click the "submit" button below to enter your responses.

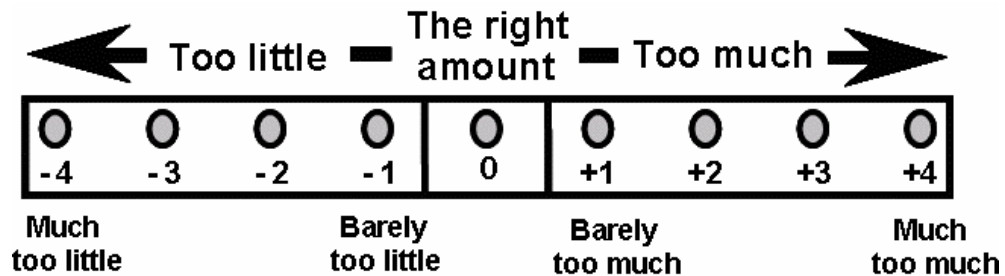
The Leadership Versatility Index - student version is derived from the Leadership Versatility Index: Copyright© 2006 Kaplan DeVries Inc. All rights reserved. The method of assessment used by the Leadership Versatility Index® is protected by U.S. Patent No. 7,121,830. Student version adapted with permission from Kaplan DeVries Inc.

In reference to items 25-54: Copyright© 2003 James M. Kouzes and Barry Z. Posner. All rights reserved. Used with permission.

APPENDIX C. LVI-S FIELD TEST ITEMS FOR PILOT STUDY

Instructions

The rating scale is different from the typical kind, where a higher score is a better score. On this scale, **the best score is a “0,” in the middle of the scale.** The premise is that there are two kinds of performance problems: when leaders emphasize something *too much* or when they put *too little* emphasis on something.



WARNING: Some people misread this scale. Please do not mistake it for the usual type where a high score is the best score.

1. Use the “**too much**” side of the scale for items that he/she takes to an extreme—what he/she **does too frequently or with too much intensity**.
2. Use the “**too little**” side for those items that he/she is deficient on—what he/she **does not do often enough or does with too little intensity**.

If you feel unable to rate a particular item because it doesn't apply or you haven't had a chance to observe the individual in that area, you may select “N/A” (not applicable). Please do not use this option more frequently than is absolutely necessary.

CONFIDENTIALITY REMINDER:

Your ratings will be anonymous. They will be averaged together with the ratings from other participants and presented as a collective average.

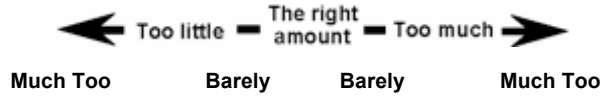
RATER CODE _____

← Too little — The right amount — Too much →

Much Too Barely Barely Much Too

| Ratee Code | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. <i>Takes Charge</i> —in control of his/her area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. Takes the initiative—eager to lead. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. Gives clear direction—tells people what to do. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. Steps in—gets actively involved when problems arise. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Decisive—makes up his/her mind quickly. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Lets people know clearly where he/she stands on issues— <i>Declares</i> him-/herself. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| ← Too little — The right amount — Too much → | | | | | | | | | | |
| 7. Tells people what is on her/his mind—forthcoming. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Defends his/her point-of-view—doesn't back down easily. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. Challenges people to do their best—Pushes hard for high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. Expects a lot—insists on high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. Direct—tells people when she/he is dissatisfied with their work. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. Holds people accountable—is firm when others do not meet his/her standards. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |

RATER CODE _____



| Ratee Code: | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 13. Empowers other people—able to step back and give them room to do their job. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 14. Gives people plenty of opportunity to show initiative. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 15. Gives people freedom to decide how to perform their duties—hands-off. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. Trusts people to handle the problems that come up in their area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 17. Participative—includes people when making decisions | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 18. Wants to know where others stand—asks for others' opinions. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | | | | | | | | | | |
| 19. Can be persuaded to change her/his mind—open to influence. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 20. Makes it safe to challenge or critique his/her thinking—welcomes dialogue and debate. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 21. Shows appreciation—goes out of his/her way to make others feel good about their contribution. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 22. Friendly—is nice and treats other people with courtesy. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 23. Sensitive—careful not to hurt the other person's feelings. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 24. Cuts people slack—is understanding when they are not able to do their duties and responsibilities. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |

I. RAs view me as being effective in meeting hall/floor needs.

| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
|-------------------------------|---------------------------------|-------------------------------|--------------------|---------------------------|-------------------------------|---------------------------------|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

II. Hall Directors and administrators view me as being effective in meeting residence hall/floor objectives.

| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
|-------------------------------|---------------------------------|-------------------------------|--------------------|---------------------------|-------------------------------|---------------------------------|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

III. I am successful at representing our residence hall/floor with other students and RAs.

| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
|-------------------------------|---------------------------------|-------------------------------|--------------------|---------------------------|-------------------------------|---------------------------------|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

IV. I am successful at representing our residence hall/floor with campus administrators.

| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
|-------------------------------|---------------------------------|-------------------------------|--------------------|---------------------------|-------------------------------|---------------------------------|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

V. I have developed a strong sense of community in this residence hall/floor.

| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
|-------------------------------|---------------------------------|-------------------------------|--------------------|---------------------------|-------------------------------|---------------------------------|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

VI. I am a positive role model as a Hall Director.

| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
|-------------------------------|---------------------------------|-------------------------------|--------------------|---------------------------|-------------------------------|---------------------------------|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

VII. When the school year is over, my supervisees and/or residents will be able to talk about the differences I made.

| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
|-------------------------------|---------------------------------|-------------------------------|--------------------|---------------------------|-------------------------------|---------------------------------|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

VIII. I am effective at getting people to behave in a responsible manner.

| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
|-------------------------------|---------------------------------|-------------------------------|--------------------|---------------------------|-------------------------------|---------------------------------|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

IX. I am able to get people to volunteer for events and responsibilities.

| | | | | | | |
|-------------------------------|---------------------------------|-------------------------------|-----------------------|---------------------------|-------------------------------|---------------------------------|
| <i>Not at all descriptive</i> | <i>Infrequently descriptive</i> | <i>Marginally descriptive</i> | <i>Descriptive</i> | <i>Fairly descriptive</i> | <i>Frequently descriptive</i> | <i>All the time descriptive</i> |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

1. Overall, to what extent do you feel the Hall Director is performing his/her job the way you would like it to be performed?

| | | | | | | |
|-----------------------|-----------------------|---------------------------------------|---------------------------------|--|-----------------------|-----------------------|
| <i>Not at all</i> | <i>A little bit</i> | <i>A little lower than moderately</i> | <i>Moderately (about 50/50)</i> | <i>A little better than moderately</i> | <i>Mostly</i> | <i>Entirely</i> |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

2. To what extent has he/she met your own expectations in his/her administrative roles and responsibilities?

| | | | | | | |
|-----------------------|-----------------------|---------------------------------------|---------------------------------|--|-----------------------|-----------------------|
| <i>Not at all</i> | <i>A little bit</i> | <i>A little lower than moderately</i> | <i>Moderately (about 50/50)</i> | <i>A little better than moderately</i> | <i>Mostly</i> | <i>Entirely</i> |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

3. If you had entirely your own way, to what extent would you change the manner in which she/he is doing the job?

| | | | | | | |
|-----------------------|-----------------------|---------------------------------------|---------------------------------|--|-----------------------|-----------------------|
| <i>Not at all</i> | <i>A little bit</i> | <i>A little lower than moderately</i> | <i>Moderately (about 50/50)</i> | <i>A little better than moderately</i> | <i>Mostly</i> | <i>Entirely</i> |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Were the instructions for this assessment clear?

What items were unclear to you?

Was the response format easy for you to understand?

How long did it take you to complete the instrument?

Do you have any suggestions that could make this instrument more clear or easier to understand?

APPENDIX D. ITEMS SUBMITTED FOR EXPERT REVIEW

(Post RA Focus Group)

LVI-Student Expert Review Feedback Form

The LVI-Student (LVI-S) is based on the Leadership Versatility Index. As such, item language needs to be adjusted to suit the college environment; the residence hall environment in particular. Two focus groups have been conducted on the LVI items, one with Housing and Residence Life experts and another with Resident Assistants (RAs). These focus groups helped ensure LVI-S content was relevant to Residence Hall leadership and the language was appropriate for college students and entry-level student affairs professionals.

Instructions: Would you help me flag items that could be problematic? Please review items and use the blank next to it to provide comments or suggested phrasing. If an item does not warrant comment, you may leave it blank. For your reference, on pages 3 & 4 I have provided a table with the original LVI items, the corresponding LVI-S items, and the rationale for the proposed changes.

| | Item | Remarks / Suggested Phrasing |
|--------------|---|------------------------------|
| Takes Charge | 1f. <i>Takes Charge</i> —in control of his/her area of responsibility. | |
| | 2f. Takes the initiative—eager to lead. | |
| | 3f. Delegates clearly—tells people what to do. | |
| | 4f. Steps in—gets actively involved when problems arise. | |
| Declares | 5f. Decisive—makes up his/her mind quickly. | |
| | 6f. Lets people know clearly where he/she stands on issues— <i>Declares</i> him-/herself. | |
| | 7f. Tells people what is on her/his mind—forthcoming. | |
| | 8f. Defends his/her point-of-view—doesn't back down easily. | |
| Pushes | 9f. Challenges people to do their best— <i>Pushes</i> hard for high performance. | |
| | 10f. Expects a lot—insists on high performance. | |
| | 11f. Direct—tells people when | |

| | | |
|----------|--|--|
| | she/he is dissatisfied with their work. | |
| | 12f. Holds people accountable—responds firmly when expectations are not met. | |
| Empowers | 1e. <i>Empowers</i> supervisees to run their units—able to step back. | |
| | 2e. Gives people plenty of opportunity to show initiative. | |
| | 3e. Gives people freedom to decide how to perform their duties—hands-off. | |
| | 4e. Trusts people to handle the problems that come up in their area of responsibility. | |
| Listens | 5e. Participative—includes people when making decisions | |
| | 6e. Wants to know where others stand—asks for others' opinions. | |
| | 7e. Can be persuaded to change her/his mind—open to influence. | |
| | 8e. Makes it safe to challenge or critique his/her thinking—welcomes debate. | |
| Supports | 9e. Shows appreciation—takes extra effort to make others feel good about their contribution. | |
| | 10e. Friendly, treats supervisees with courtesy. | |
| | 11e. Sensitive—careful not to hurt the other person's feelings. | |
| | 12e. Cuts people slack—considerate of others' circumstances. | |

APPENDIX E. FIRST REVISED SET OF ITEMS

(Post Administrator Focus Group, Pre-RA Focus Group)

| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. <i>Takes Charge</i> —in control of his/her area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. Takes the initiative—eager to lead. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. Sets clear expectations—tells people what to do. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. Steps in—gets personally involved when problems arise. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Decisive—makes up his/her mind quickly. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Clearly states where he/she stands on issues— <i>Declares</i> him-/herself. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | | | | | | | | | | |
| 7. Tells people what is on her/his mind—forthcoming. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Defends his/her position—doesn't back down easily. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. Challenges people to do their best— <i>Pushes</i> hard for high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. Expects a lot—insists on high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. Direct—tells people when she/he is dissatisfied with their work. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. Holds people accountable—responds firmly when expectations are not met. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |

| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 13. <i>Empowers</i> supervisees to run their units—able to step back. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 14. Gives people plenty of opportunity to show initiative. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 15. Gives people freedom to decide how to perform their duties—hands-off. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. Trusts people to handle the problems that come up in their area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 17. Participative—includes others when making decisions | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 18. Wants to know where others stand—asks for others' opinions. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | | | | | | | | | | |
| 19. Can be persuaded to change her/his mind—open to influence. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 20. Makes it safe to challenge or critique for people his/her thinking—welcomes debate. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 21. Shows appreciation—takes extra effort to make others feel good about their contribution. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 22. Friendly treats supervisees with courtesy. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 23. Sensitive—careful not to hurt the other person's feelings. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 24. Cuts people slack—considerate of others' circumstances. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |

APPENDIX F. INITIAL SET OF PROPOSED ITEMS

(Prior to Administrator Focus Group Feedback)

| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. <i>Takes Charge</i> —in control of his/her area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. Takes the initiative—eager to lead. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. Sets clear expectations—tells people what to do. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. Steps in—gets personally involved when problems arise. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Decisive—makes up his/her mind quickly. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Clearly states where he/she stands on issues— <i>Declares</i> him-/herself. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | | | | | | | | | | |
| 7. Tells people what is on her/his mind—forthcoming. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Defends his/her position—doesn't back down easily. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. <i>Pushes</i> others hard. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. Expects a lot—insists on high performance. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. Direct—tells people when she/he is dissatisfied with their work. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. Holds people accountable—responds firmly when expectations are not met. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |

| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 13. <i>Empowers</i> people to decide how to do their work—able to let go. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 14. Gives people plenty of room to show initiative. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 15. Gives people freedom to decide how to perform their duties—hands-off. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. Trusts people to handle the problems that come up in their area of responsibility. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 17. Participative—includes others when making decisions | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 18. Wants to know where others stand—asks for others' opinions. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | | | | | | | | | | |
| 19. Can be persuaded to change her/his mind—open to influence. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 20. Makes it easy for people to challenge his/her thinking. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 21. Shows appreciation—takes extra effort to make others feel good about their contribution. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 22. Nice to people, treats them well. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 23. Sensitive—careful not to hurt the other person's feelings. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 24. Cuts people slack—considerate of others' circumstances. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | N/A |

APPENDIX G. PERMISSION FOR LEADERSHIP PRACTICES INVENTORY

KOUZES POSNER INTERNATIONAL

15419 Banyan Lane
Monte Sereno, California 95030 USA
FAX: (408) 354-9170

September 24, 2008

Mr. Preston Yarborough
4027 Quartergate Drive
High Point, North Carolina 27265

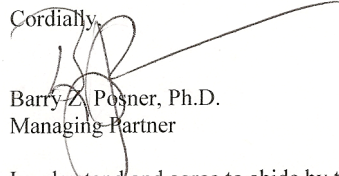
Dear Preston:

Thank you for your request to use the Leadership Practices Inventory (LPI) in your dissertation. We are willing to allow you to *reproduce* the instrument in written form as outlined in your request, at no charge, with the following understandings:

- (1) That the LPI is used only for research purposes and is not sold or used in conjunction with any compensated management development activities;
- (2) That copyright of the LPI, or any derivation of the instrument, is retained by the authors, and that the following copyright statement is included on all copies of the instrument: "Copyright © 2003 James M. Kouzes and Barry Z. Posner. All rights reserved. Used with permission.";
- (3) That one (1) **electronic** copy of your dissertation and one (1) copy of **all** papers, reports, articles, and the like which make use of the LPI data be sent **promptly** to our attention; and,
- (4) That you agree to allow us to include an abstract of your study and any other published papers utilizing the LPI on our various websites.

If the terms outlined above are acceptable, would you indicate so by signing one (1) copy of this letter and returning it to us. Best wishes for every success with your research project.

Cordially,


Barry Z. Posner, Ph.D.
Managing Partner

I understand and agree to abide by these conditions:

(Signed)  Date: 2/28/09

APPENDIX H. PERMISSION FOR LEADERSHIP VERSATILITY INDEX

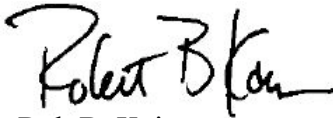
Dear Preston:

Thank you for your interest in the versatile leader model and the Leadership Versatility Index (LVI). You wish to modify the LVI for use with college students and entry-level professionals. We will provide you permission to modify the LVI for your dissertation with the following understandings:

- (1) That your modified instrument is used only for research purposes and is not sold or used in conjunction with any compensated management development activities;
- (2) That copyright of the Too Little / Too Much response scale, the LVI, or any derivation of the instrument, is retained by the authors, and that the following copyright statement is included on all copies of the instrument: "Copyright © 2006 Kaplan DeVries, Inc. All rights reserved. Used with permission.";
- (3) That one (1) electronic copy of your dissertation and one (1) copy of all papers, reports, articles, and the like which make use of the LVI data be sent promptly to our attention; and,
- (4) That you agree to allow us to include an abstract of your study and any other published papers utilizing the LVI on our various websites.

If the terms outlined above are acceptable would you indicate so by signing one (1) copy of this letter and returning it to us. Best wishes for every success with your research project.

Cordially,



Rob B. Kaiser
Partner

I understand and agree to abide by these conditions:

(signed) _____ Date: _____

APPENDIX I. INFORMED CONSENT, FOCUS GROUP PARTICIPANT

FOCUS GROUP PARTICIPANT

UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

Section 1.01

CONSENT TO ACT AS A HUMAN

PARTICIPANT

Project Title: Development and Validation of the Leadership Versatility Index for Students

Project Director: James Preston Yarborough

Participant's Name: _____

What is the study about?

This research study is focused on the development a multi-rater (360-degree) leadership assessment instrument (Leadership Versatility Index for Students—LVI-S) for students and entry-level professionals working in Housing and Residence Life. This instrument explores leadership versatility, or a leader's ability to effectively use a wide range of leadership behaviors. Few multi-rater leadership instruments have been developed specifically for collegiate audiences. The LVI-S will help address that need.

Why are you asking me?

You have been selected for participation in this focus group because you are either considered an expert on content relevant to this instrument or because you are currently serving as a Hall Director or Resident Advisor and have direct knowledge of the roles and responsibilities of these positions.

What will you ask me to do if I agree to be in the study?

Focus group participants will provide an initial review of item content and clarity. You will be asked to provide oral and/or written feedback on a proposed set of items for this study. Focus groups will be conducted with Residence Hall administrators, Hall Directors, and Resident Advisors. The focus groups should last approximately 60 minutes. None of these procedures are expected to involve stress, pain, or produce any other unpleasant reaction.

Is there any audio/video recording?

N/A

What are the dangers to me?

There are no dangers associated with participating in this study.

Some potential participants in this study are students at UNC-Greensboro or are staff or faculty at UNC-Greensboro. Participation in this study is strictly voluntary. If you choose to participate, the information you provide will be strictly confidential and will not impact your relationship with

Preston Yarborough or the Office of Leadership and Service-Learning. Should you choose not to participate this will in no way affect your relationship or standing with Preston Yarborough or his employer, the Office of Leadership and Service-Learning at UNC-Greensboro.

Focus group participants have multiple ways to submit data. You may participate in the discussion, in which case your perspectives and insights will be shared with fellow focus-group participants. Should you prefer to submit candid comments, you may use a feedback form to communicate these observations. Your feedback form will not be shared with other participants and unless you orally share the contents of this form during the discussion portion of the focus group, fellow participants are unlikely to know what you have written. I cannot guarantee that a member will not try to look at your paper, however, and I cannot control what one focus group member decides to share with another member.

If you have any concerns about your rights or how you are being treated please contact Eric Allen in the Office of Research and Compliance at UNCG at (336) 256-1482. Questions about this project or your benefits or risks associated with being in this study can be answered by Preston Yarborough who may be contacted at (336) 803-2149 or by e-mail at jpyarbor@uncg.edu.

Are there any benefits to me for taking part in this research study?

Participants in this study are interested in helping improve student /entry-level professional leadership skills. These participants might derive satisfaction from contributing to a process that can advance our understanding of leadership development.

Are there any benefits to society as a result of me taking part in this research?

Leadership research has empirically demonstrated the need and value for multi-rater feedback data. Multi-rater feedback data is more reliable and has significantly greater predictive validity than self-reported leadership data. Very few multi-rater leadership instruments are appropriate for use in college settings. This study will produce a new multi-rater instrument that can measure participants' effective (or ineffective) use of leadership skills. This information may be used to inform leadership development training and to evaluate results of training initiatives.

Will I get paid for being in the study? Will it cost me anything?

There are no costs to you or payments made for participating in this study.

How will you keep my information confidential?

Consent forms for all participants will be maintained in a locked filing cabinet in the researcher's place of residence. These forms will be filed and separately from the data collected in the study, but both will be stored in separate drawers of this filing cabinet. Data collected will include feedback forms from the expert review process, facilitator notes from the focus group sessions, and any written comments or observations received from participants in the focus groups.

During the pilot study several measures have been enacted to preserve your confidentiality. First, your identity will not be revealed. Focus group participants will complete a paper and pencil form that will be assigned an ID code that prevents documents from having your name on them. This ID code will be unique to you. There is a remote possibility that someone with the code could identify you as a participant. To minimize this risk several procedures will be followed. Only the researcher will possess a key to this code and only the researcher can access the key. The code will be kept in a password-protected file on the researcher's personal computer..

The consent forms and data from this study will be kept for 3 years. After this duration, the documents will be destroyed through the use of a paper shredder.

What if I want to leave the study?

You have the right to refuse to participate or to withdraw at any time, without penalty. If you do withdraw, it will not affect you in any way. If you choose to withdraw, you may request that any of your data which has been collected be destroyed unless it is in a de-identifiable state.

What about new information/changes in the study?

If significant new information relating to the study becomes available which may relate to your willingness to continue to participate, this information will be provided to you.

Voluntary Consent by Participant:

By signing this consent form you are agreeing that you read, or it has been read to you, and you fully understand the contents of this document and are openly willing consent to take part in this study. All of your questions concerning this study have been answered. By signing this form, you are agreeing that you are 18 years of age or older and are agreeing to participate, or have the individual specified above as a participant participate, in this study described to you by Preston Yarborough.

Signature: _____ Date: _____

APPENDIX J. INFORMED CONSENT, PILOT STUDY HALL DIRECTOR

UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

Section 1.02

CONSENT TO ACT AS A HUMAN

PARTICIPANT: LONG FORM

Project Title: Development and Validation of the Leadership Versatility Index for Students

Project Director: James Preston Yarborough

Participant's Name: _____

What is the study about?

This research study is focused on the development a multi-rater (360-degree) leadership assessment instrument (Leadership Versatility Index for Students—LVI-S) for students and entry-level professionals working in Housing and Residence Life. This instrument explores leadership versatility, or a leader's ability to effectively use a wide range of leadership behaviors. Few multi-rater leadership instruments have been developed specifically for collegiate audiences. The LVI-S will help address that need.

Why are you asking me?

You have been selected for participation in this pilot phase because you are serving as a Hall Director or Resident Advisor and have direct knowledge of the roles and responsibilities of these positions.

What will you ask me to do if I agree to be in the study?

You will participate in the administration of a pilot version of the LVI-S. Participants in this trial study will be Hall Directors and the Resident Advisors whom they supervise. In a multi-rater instrument, the focal leader (Hall Director) provides self-ratings of his/her leadership behaviors on a hand-written, self-report form. Completion of this form is not expected to exceed 30 minutes. Observers (Resident Advisors) will also provide ratings of you on a hand-written observer-report form. Once instruments are completed, they will be inserted into envelopes and returned to the investigator for scoring and data analysis.

Is there any audio/video recording?

N/A

What are the dangers to me?

Some potential participants in this study are students at UNC-Greensboro or are staff or faculty at UNC-Greensboro. Participation in this study is strictly voluntary. If you choose to participate, the information you provide will be strictly confidential and will not impact your relationship with Preston Yarborough or the Office of Leadership and Service-Learning. Should you choose not to participate, this decision will in no way affect your relationship or standing with Preston Yarborough or his employer, the Office of Leadership and Service-Learning at UNC-Greensboro.

Hall Directors (the focal leaders) participating in the field test are being rated on their leadership characteristics. Having someone rate your leadership characteristics could be perceived as a somewhat stressful event. Several measures have been implemented to ease this concern. Your participation in this pilot study will remain confidential and any data resulting from this study will be reported anonymously. In other words, when data is reported, you will not be identified nor will any of your observers. Furthermore, multiple Hall Directors will be participating in the pilot study. Even if someone knew you participated in the study, it would be difficult for this person to look at the results and figure out which set of data pertained to you.

If you have any concerns about your rights or how you are being treated please contact Eric Allen in the Office of Research and Compliance at UNCG at (336) 256-1482. Questions about this project or your benefits or risks associated with being in this study can be answered by Preston Yarborough who may be contacted at (336) 803-2149 or by e-mail at jpyarbor@uncg.edu.

Are there any benefits to me for taking part in this research study?

Participants in this study are interested in helping improve student /entry-level professional leadership skills. These participants might derive satisfaction from contributing to a process that can advance our understanding of leadership development.

Are there any benefits to society as a result of me taking part in this research?

Leadership research has empirically demonstrated the need and value for multi-rater feedback data. Multi-rater feedback data is more reliable and has significantly greater predictive validity than self-reported leadership data. Very few multi-rater leadership instruments are appropriate for use in college settings. This study will produce a new multi-rater instrument that can measure participants' effective (or ineffective) use of leadership skills. This information may be used to inform leadership development training and to evaluate results of training initiatives.

Will I get paid for being in the study? Will it cost me anything?

There are no costs to you or payments made for participating in this study.

How will you keep my information confidential?

Consent forms for all participants will be maintained in a locked filing cabinet in the researcher's place of residence. These forms will be filed and separately from the data collected in the study, but both will be stored in separate drawers of this filing cabinet. Data collected from Phase I will include feedback forms from the expert review process, facilitator notes from the focus group sessions, and any written comments or observations received from participants in the focus groups.

During the pilot study several measures have been enacted to preserve your confidentiality. First, your identity will not be revealed. Pilot study participants will complete a paper and pencil form that will be assigned an ID code that prevents documents from having your name on them. The focal leader will be provided an ID number; observers for that leader will receive ID numbers that are a subset of the leader's ID. Through this process, leaders' data and their respective observers' data can be matched without having to use names or other outwardly recognizable identification. This ID code will be unique to you. There is a remote possibility that someone with the code could identify you as a participant. To minimize this risk several procedures will be followed. Only the researcher will possess a key to this code and only the researcher can access the key. The code will be kept in a password-protected file on the researcher's personal computer. To minimize the risk of someone looking at the reported data and attempting to deduce who the participants were, more than one Hall Director (focal leader) will be participating in the pilot study. Having more than one focal leader reduces the likelihood of someone looking at the reported outcome data and deducing the participants of the study.

The consent forms and data from this study will be kept for 3 years. After this duration, the documents will be destroyed through the use of a paper shredder.

What if I want to leave the study?

You have the right to refuse to participate or to withdraw at any time, without penalty. If you do withdraw, it will not affect you in any way. If you choose to withdraw, you may request that any of your data which has been collected be destroyed unless it is in a de-identifiable state.

What about new information/changes in the study?

If significant new information relating to the study becomes available which may relate to your willingness to continue to participate, this information will be provided to you.

Voluntary Consent by Participant:

By signing this consent form you are agreeing that you read, or it has been read to you, and you fully understand the contents of this document and are openly willing consent to take part in this study. All of your questions concerning this study have been answered. By signing this form, you are agreeing that you are 18 years of age or older and are agreeing to participate, or have the individual specified above as a participant participate, in this study described to you by Preston Yarborough.

Signature: _____ Date: _____

APPENDIX K. INFORMED CONSENT, EXPERT REVIEW

UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

Section 1.03

CONSENT TO ACT AS A HUMAN

PARTICIPANT: LONG FORM

Project Title: Development and Validation of the Leadership Versatility Index for Students

Project Director: James Preston Yarborough

Participant's Name: _____

What is the study about?

This research study is focused on the development a multi-rater (360-degree) leadership assessment instrument (Leadership Versatility Index for Students—LVI-S) for students and entry-level professionals working in Housing and Residence Life. This instrument explores leadership versatility, or a leader's ability to effectively use a wide range of leadership behaviors. Few multi-rater leadership instruments have been developed specifically for collegiate audiences. The LVI-S will help address that need.

Why are you asking me?

You have been selected for participation in this pilot phase because you are either considered an expert on content relevant to this instrument or are an expert on assessment instrument design.

What will you ask me to do if I agree to be in the study?

An expert review panel will review items to ensure developmentally appropriate language and that items appear appropriately designed for the research context. Your feedback will be provided in written form on documents provided for you by the researcher. Once you have completed your review of the items, you are asked to return them to the researcher in the provided, pre-addressed envelope. While individual review times may vary, it is anticipated the item review process should take approximately 30 minutes.

Is there any audio/video recording?

N/A

What are the dangers to me?

There are no dangers associated with participating in this study.

Some potential participants in this study are students at UNC-Greensboro or are staff or faculty at UNC-Greensboro. Participation in this study is strictly voluntary. If you choose to participate, the information you provide will be strictly confidential and will not impact your relationship with Preston Yarborough or the Office of Leadership and Service-Learning. Should you choose not to

participate this will in no way affect your relationship or standing with Preston Yarborough or his employer, the Office of Leadership and Service-Learning at UNC-Greensboro.

Hall Directors (the focal leaders) participating in the field test are being rated on their leadership ability. Having someone rate your leadership characteristics could be perceived as a somewhat stressful event; similarly, rating one's supervisor could also be perceived as a somewhat stressful event. The results from this pilot study will remain confidential. No participants in the study will receive information about how other participants responded. All results reported as a product of this study will be presented in aggregate form.

If you have any concerns about your rights or how you are being treated please contact Eric Allen in the Office of Research and Compliance at UNCG at (336) 256-1482. Questions about this project or your benefits or risks associated with being in this study can be answered by Preston Yarborough who may be contacted at (336) 803-2149 or by e-mail at jpyarbor@uncg.edu.

Are there any benefits to me for taking part in this research study?

Participants in this study are interested in helping improve student /entry-level professional leadership skills. These participants might derive satisfaction from contributing to a process that can advance our understanding of leadership development.

Are there any benefits to society as a result of me taking part in this research?

Leadership research has empirically demonstrated the need and value for multi-rater feedback data. Multi-rater feedback data is more reliable and has significantly greater predictive validity than self-reported leadership data. Very few multi-rater leadership instruments are appropriate for use in college settings. This study will produce a new multi-rater instrument that can measure participants' effective (or ineffective) use of leadership skills. This information may be used to inform leadership development training and to evaluate results of training initiatives.

Will I get paid for being in the study? Will it cost me anything?

There are no costs to you or payments made for participating in this study.

How will you keep my information confidential?

Consent forms for all participants will be maintained in a locked filing cabinet in the researcher's place of residence. These forms will be filed and separately from the data collected in the study, but both will be stored in separate drawers of this filing cabinet. Data collected will include feedback forms from the expert review process, facilitator notes from the focus group sessions, and any written comments or observations received from participants in the focus groups.

During the expert review process, several measures have been enacted to preserve your confidentiality. Expert review participants will complete a paper and pencil form that will be assigned an ID code that prevents documents from having your name on them. This ID code will be unique to you. There is a remote possibility that someone with the code could identify you as a participant. To minimize this risk several procedures will be followed. Only the researcher will possess a key to this code and only the researcher can access the key. The code will be kept in a password-protected file on the researcher's personal computer.

The consent forms and data from this study will be kept for 3 years. After this duration, the documents will be destroyed through the use of a paper shredder.

What if I want to leave the study?

You have the right to refuse to participate or to withdraw at any time, without penalty. If you do withdraw, it will not affect you in any way. If you choose to withdraw, you may request that any of your data which has been collected be destroyed unless it is in a de-identifiable state.

What about new information/changes in the study?

If significant new information relating to the study becomes available which may relate to your willingness to continue to participate, this information will be provided to you.

Voluntary Consent by Participant:

By signing this consent form you are agreeing that you read, or it has been read to you, and you fully understand the contents of this document and are openly willing consent to take part in this study. All of your questions concerning this study have been answered. By signing this form, you are agreeing that you are 18 years of age or older and are agreeing to participate, or have the individual specified above as a participant participate, in this study described to you by Preston Yarborough.

Signature: _____ Date: _____

APPENDIX L. PILOT STUDY SCRIPT

- You are being asked if you want to be in a research study. We are trying to find out about Resident Advisors' perspectives on their Hall Director's leadership characteristics. The purpose of this study is to develop a new, multi-rater leadership instrument.
- You have been selected for participation in this study because you are either considered an expert on content relevant to this instrument or because you are currently serving as a Hall Director or Resident Advisor and have direct knowledge of the roles and responsibilities of these positions.
- This discussion and the piece of paper (short form) given to you will tell you about the study to help you decide if you want to be part of the study
- You will be asked to participate in a (focus group; or expert review process; or field test of the pilot version of the instrument). Your time commitment should not exceed (one hour for focus group participation; 30 minutes for expert review participation; or 30 minutes for field test participation). There are no costs for participating in this pilot study.
- There are no payments made for participating in this study

The benefits to (you and or society) being in this study include helping improve our ability to measure student /entry-level professional leadership skills. Leadership research has demonstrated the need and value for multi-rater feedback data. Multi-rater feedback data helps us understand how followers perceive their leaders. Since leadership is a social behavior, including followers' perspective of leaders is an important, but frequently overlooked aspect of student leadership research. This study will produce a new multi-rater instrument that can measure participants' effective (or ineffective) use of leadership skills. This information may be used to inform leadership development training and to evaluate results of training initiatives.

Other than a slight breach of confidentiality, which is always present when identifiable data is collected, several aspects of this study pose no anticipated risks. Those who participate in the field test portion of the study, might feel some element of risk. Hall Directors (the focal leaders) participating in the field test are being rated on their leadership ability. Having someone rate your leadership characteristics could be perceived as a somewhat stressful event. Similarly, rating one's supervisor could also be perceived as a somewhat stressful event. The results from this pilot study will remain confidential. No participants in the study will receive information about how other participants responded. All results reported as a product of this study will be presented in aggregate form.

Your privacy will be protected. Consent forms for all participants will be maintained in a locked filing cabinet at my place of residence.

Data collected from Phase I will be maintained in a locked filing cabinet. This data includes response forms for the expert review process, facilitator notes from the focus group sessions, and any written comments or observations received from participants in the focus groups.

Phase II field study participants will complete an electronic form that is coded in such a way that participants are not required to use their names. The focal leader will be provided an ID number;

observers for that leader will receive ID numbers that are a subset of the leader's ID. Through this process, leaders' and their respective observers' data can be matched without having to use names or other outwardly recognizable identification. Electronic data will be password protected and stored in a server with external firewall protection. Absolute confidentiality of data provided through the Internet cannot be guaranteed due to the limited protection of Internet access. Please be sure to close your browser when finished so no one will be able to see what you have been doing.

- All information obtained in this study is strictly confidential unless disclosure is required by law.
- You should ask any questions you have before making up your mind. You can think about it and talk to your family or friends before you decide if you want to be in the study
- If you decide you want to be in the study you will need to sign the piece of paper (short form) given to you earlier. A family member, friend, or someone next to you will also need to sign this piece of paper as the witness.
- If you decide you do not want to be in the study later you are free to leave when ever you like without penalty or unfair treatment.

APPENDIX M. PILOT STUDY RECRUITMENT E-MAIL

From: Curtis Erwin CGERWIN/facultystaff/uncg
To: Preston Yarborough JPYARBOR/facultystaff/uncg@uncg

Date: Friday, January 16, 2009 08:10AM
Subject: Re: Pilot Study Groups

History: This message has been replied to.

Wanted you to know I have received your request. This may take a little work as we are extremely under-staffed and it is difficult to keep asking for persons to do additional items that go beyond their required daily tasks. I will bring up next week, I am sure we will get some volunteers but want to ask personally, and not thorough an e-mail. RA participation should be easier to achieve.

Curtis G. Erwin
Associate Director for Residence Life
University of North Carolina - Greensboro
336.334.5636 (office), 336.334.5680 (fax)
Preston Yarborough JPYARBOR/facultystaff/uncg

Preston Yarborough
JPYARBOR/facultystaff/uncg

01/15/2009 11:40 AM

ToCurtis Erwin
CGERWIN/facultystaff/uncg@uncg
cc
SubjectPilot Study Groups

Curt,

To finalize the instrument development process, I'll need to consult with content experts. You and your gang are the experts, so I thought I would provide a sketch of the next steps. I'll do the coordinating, but if you could suggest some personnel that would be a good fit with the following processes, I would be indebted to you (again).

After the IRB committee reviews my paperwork, I need to convene two focus groups and identify a couple of Hall Directors and RAs to take a trial version of the instrument. My goal is to propose my study in mid-February-early March. Then collect data for the main study shortly thereafter.

Step One: Focus Group I--HRL Professional Administrators, related experts, and Hall Director(s); 60 minutes

Number : I'd like to have 2-3 administrators, hopefully Deb Taub or someone from Higher Ed, and about 2 Hall Directors

Purpose : Brief you on underlying premise of the instrument, present rough draft of items, receive feedback on the items (are they developmentally appropriate for students, are they contextually relevant for an HRL setting?)

Step Two: Focus Group II--Resident Advisors; 60 minutes

Numbers : 5-8 Resident Advisors, preferably a diverse group

Purpose : Review item language and strive to ensure final item set imparts the intended images.

Step Three: Trial Run of Assessment; 60 minutes

APPENDIX N. MAIN STUDY HALL DIRECTOR RECRUITMENT

Initial Hall Director Recruitment Letter—Dept of Residence Life

(To be sent by HRL Department)

Dear Hall Director Team, [Or insert an introduction appropriate for your group]

Preston Yarborough, a doctoral student in the Department of Counseling and Educational Development at The University of North Carolina at Greensboro, has approached our department for assistance with his dissertation study on residence hall leadership. It explores how supervisors perceive their personal leadership characteristics as well as how Resident Advisors perceive the leadership characteristics of their supervisor. The study also explores how these characteristics relate to leader effectiveness.

This information will be collected in an on-line survey that will take approximately 15-20 minutes to complete. All participant data will be kept confidential and any reported data will be presented in aggregate form (in other words, no individual-level data will be reported). Should you have further questions about the study, questions about confidentiality, or about the type of feedback available (individual vs. group), please contact Preston at jpyarbor@uncg.edu.

Participation in this study can help us better understand the nature of effective leadership in residence halls. You will help validate a new assessment tool that can enhance the leadership development and training of residence hall supervisors and RAs. Upon completion of the study, a mini-colloquium summarizing the findings will be presented to our Department. Hall Directors interested in learning more about how these findings can be of service to their leadership development will be invited to this event.

Participants in this study will have an opportunity to win an i-Pod shuffle. The winner of the shuffle will be determined by random drawing.

Your participation in this study is voluntary and in no way affects your standing with this department. We will not know if you participate or if you do not. If you choose to participate but then wish to opt out of the study, you may do so at any time without penalty.

To participate in the study, please click the link below or paste it in your browser window.

[www.versatileleadershipstudy.com]

Sincerely,

[Insert Your Name]

APPENDIX O. MAIN STUDY FOLLOW-UP HALL DIRECTOR RECRUITMENT

(Sent directly by Preston Yarborough)

E-mail Subject: Residence Hall Leadership Study

My name is Preston Yarborough and I am a doctoral student at The University of North Carolina at Greensboro in the Department of Counseling and Educational Development. Your Department has expressed interest in my dissertation study and has offered permission for me to ask if you would be a volunteer participant.

I have developed an assessment that measures the leadership characteristics of residence hall supervisors. My study explores how supervisors perceive their personal leadership characteristics as well as how Resident Advisors perceive the leadership characteristics of their supervisor. It also studies how these characteristics relate to leader effectiveness.

You will be asked to complete a 15 -20 minute on-line survey on your leadership characteristics. The RAs you supervise will complete a similar survey, but it will ask them to rate your leadership characteristics.

Participants may feel some risk at having other people rate their leadership characteristics and their effectiveness. All participant data will be kept confidential and any reported data will be presented in aggregate form (in other words, no individual-level data will be reported). Your participation can help us better understand the nature of effective leadership in residence halls. You will be helping to validate a new assessment tool that can enhance the leadership development and training of residence hall supervisors and RAs.

Participants in this study will have an opportunity to win an i-Pod shuffle. The winner of the shuffle will be determined by a random drawing. When the study is completed and data are analyzed, a mini-colloquia that summarizes the findings will be presented to your Department. Participating Hall Directors interested in learning more about how these findings can be of service to their leadership development will be invited to the colloquia.

I want to emphasize that your participation is voluntary and does not affect your standing with your department. If you choose to participate but then wish to opt out of the study, you may do so at any time without penalty.

Sincerely,

Preston Yarborough
Doctoral Candidate
Department of Counseling and Educational Development
The University of North Carolina at Greensboro

APPENDIX P. MAIN STUDY RA RECRUITMENT LETTER—HALL DIRECTOR

(To be sent by Hall Director upon completion of Survey)

Dear Resident Advisor Team, [Or insert an introduction appropriate for your group]

I have agreed to participate in a dissertation study on residence hall leadership. This study will help us understand residence hall leadership more effectively and will help produce a student-focused leadership assessment that measures how leaders balance different types of leadership behaviors.

Part of this study involves my staff (you) providing your opinion of my leadership style. This information will be collected in a survey distributed by Preston Yarborough, the author of this study. Should you choose to participate, please know your ratings will remain anonymous and feel free to answer honestly. Should you have further questions, please contact Preston at jpyarbor@uncg.edu.

Where possible, Preston would like to share his findings with our department in order to provide insight into our leadership styles and leadership development opportunities. Any information I receive as a part of this study will be in aggregate form. This means your responses will be mixed with the responses of at least 5 other people. Participation is completely voluntary. You may choose not to participate without penalty.

If you choose to participate, you have the option of entering yourself into a drawing for an i-Pod Shuffle. You will also help us learn more about effective leadership and will be contributing toward the development of an innovative leadership evaluation tool.

To participate in the study, please click the link below or paste it into your browser. The survey form should take about 15-20 minutes to complete.

[www.versatileleadershipstudy.com]

Sincerely,

[Insert Your Name]

APPENDIX Q. MAIN STUDY FOLLOW-UP RA RECRUITMENT LETTER

(Sent directly by Preston Yarborough)

E-mail Subject: Residence Hall Leadership Study

My name is Preston Yarborough and I am a doctoral student at The University of North Carolina at Greensboro in the Department of Counseling and Educational Development. Your Department has expressed interest in my dissertation study and has offered permission for me to ask if you would be a volunteer participant.

I have developed an assessment that measures the leadership characteristics of residence hall supervisors. My study explores how supervisors perceive their personal leadership characteristics as well as how Resident Advisors perceive the leadership characteristics of their supervisor. It also studies how these characteristics relate to leader effectiveness.

You will be asked to complete a 15 -20 minute on-line survey on your supervisor's leadership characteristics.

Participants may feel some risk rating the leadership characteristics and effectiveness of their supervisor. Your supervisor will not know which RAs have provided data because all participant data will be kept confidential. Furthermore, any data reported from this study will be presented in aggregate form; your ratings of your supervisor will be combined with others' ratings. Your participation can help us better understand the nature of effective leadership in residence halls. You will be helping to validate a new assessment tool that can enhance the leadership development and training of residence hall supervisors and RAs.

Participants in this study will have an opportunity to win an i-Pod shuffle. The winner of the shuffle will be determined by a random drawing. When the study is completed and data are analyzed, a mini-colloquia that summarizes the findings that will be presented to your Department.

I want to emphasize that your participation is voluntary and does not affect your standing with your department. If you choose to participate but then wish to opt out of the study, you may do so at any time without penalty.

Sincerely,

Preston Yarborough
Doctoral Candidate
Department of Counseling and Educational Development
The University of North Carolina at Greensboro

APPENDIX R: HALL DIRECTOR SUPERVISOR RECRUITMENT LETTER

E-mail Subject: Residence Hall Leadership Study

My name is Preston Yarborough and I am a doctoral student at The University of North Carolina at Greensboro in the Department of Counseling and Educational Development. Your Department has expressed interest in my dissertation study and has offered permission for me to ask if you would be a volunteer participant.

I have developed an assessment that measures the leadership characteristics of residence hall supervisors. My study explores how supervisors perceive their personal leadership characteristics as well as how Resident Advisors perceive the leadership characteristics of their supervisor. It also studies how these characteristics relate to leader effectiveness.

You will be asked to complete a very brief (3-item) multiple-choice questionnaire on the effectiveness of the Hall Directors you supervise. While you will need to complete a set of questions for each Hall Director you supervise, each set of questions should take less than 2 minutes to complete.

These ratings help to correct for a common form of bias in leadership research called same source bias. When behavioral ratings are collected from a source and then effectiveness ratings are taken from that same source, then the validity is limited by same source bias (Podsakoff & Organ, 1986). Collecting effectiveness ratings from the Hall Director's supervisor and collecting behavioral ratings from RAs reporting to the Hall Director helps control for same-source bias.

Participants may feel some risk at rating a person's effectiveness for the purposes of a study. All participant data will be kept confidential and any reported data will be presented in aggregate form (in other words, no individual-level data will be reported). Your participation can help us better understand the nature of effective leadership in residence halls. You will be helping to validate a new assessment tool that can enhance the leadership development and training of residence hall supervisors and RAs.

When the study is completed and data are analyzed, a mini-colloquia that summarizes the findings will be presented to your Department. Participants interested in learning more about how these findings can be of service to their leadership development and the leadership development of the department will be invited to the colloquia.

I want to emphasize that your participation is voluntary and does not affect your standing with me or your department. If you choose to participate but then wish to opt out of the study, you may do so at any time without penalty.

Sincerely,

Preston Yarborough
Doctoral Candidate
Department of Counseling and Educational Development
The University of North Carolina at Greensboro

APPENDIX S. INSTITUTIONAL LETTER OF SUPPORT EXAMPLE

Dr. James Benshoff
Department of Counseling and Educational Development
UNCG
P.O. Box 26170
Greensboro, NC 27402-6170

Dear Dr. Benshoff,

Pat Sample, Housing and Residence Life is committed to the development of resources and tools that increase our ability to teach and measure leadership skills in the housing and residence life context. We support Mr. Yarborough's efforts to recruit Hall Directors, Resident Advisors, and administrators to complete the surveys associated with his study.

This letter is to provide permission for Preston Yarborough to recruit participants, disperse his survey, analyze the data, and present the findings for his dissertation. My staff and I are committed to working with the researcher in his data collection effort. I understand the project proposal will be reviewed and approved by UNCG Institutional Review Board for Research Involving Human Participants prior to data collection.

If you need further information in support of this project please contact me at 336.555.5555 or psample@wxyz.edu.

Sincerely yours,

Pat Sample
Associate Director for Housing & Residence Life

APPENDIX T. MAIN STUDY INFORMED CONSENT, HALL DIRECTOR

UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

Section 1.04

CONSENT TO ACT AS A HUMAN

PARTICIPANT: LONG FORM

Project Title: Development and Validation of the Leadership Versatility Index for Students

Project Director: James Preston Yarborough

What is the study about?

This research study is focused on the development a multi-rater (360-degree) leadership assessment instrument (Leadership Versatility Index for Students—LVI-S) for students and entry-level professionals working in Housing and Residence Life. This instrument explores leadership versatility, or a leader’s ability to effectively use a wide range of leadership behaviors. Few multi-rater leadership instruments have been developed specifically for collegiate audiences. The LVI-S will help address that need.

Why are you asking me?

You have been selected for participation in this pilot phase because you are serving as a Hall Director or Resident Advisor and have direct knowledge of the roles and responsibilities of these positions.

What will you ask me to do if I agree to be in the study?

You will participate in the initial administration of the LVI-S. Participants in this study will be Hall Directors, the Resident Advisors whom they supervise, and the Hall Directors’ supervisors. In a multi-rater instrument, the focal leader (Hall Director) provides self-ratings of his/her leadership behaviors on an electronic self-report survey form. Completion of this form should take approximately 20 minutes though you may take as long as you like. Observers (Resident Advisors) will provide ratings of you on a similar observer-report form. Your supervisor will provide a brief assessment of your overall effectiveness. Once you press the “submit” button at the end of the survey, the data will be transferred into a password-protected database.

Is there any audio/video recording?

N/A

What are the dangers to me?

Some potential participants in this study are students at UNC-Greensboro or are staff or faculty at UNC-Greensboro. Participation in this study is strictly voluntary. If you choose to participate, the information you provide will be strictly confidential and will not impact your relationship with Preston Yarborough, the Office of Leadership and Service-Learning, or your standing with your current department of employment. Should you choose not to participate, this decision will in no way affect your relationship or standing with Preston Yarborough or his employer, the Office of Leadership and Service-Learning at UNC-Greensboro. It will not affect your relationship with your current department of employment.

Hall Directors (the focal leaders) participating in this study are being rated on their leadership characteristics. Having someone rate your leadership characteristics and your effectiveness in

your role could be perceived as a somewhat stressful event. Several measures have been implemented to ease this concern. Your participation in this study will remain confidential and any data resulting from this study will be reported anonymously and in aggregate form. When data is reported, you will not be identified nor will any of your observers. Furthermore, many Hall Directors will be participating in the study. Even if someone knew you participated, it would be difficult for this person to look at the results and figure out how your individual ratings contributed to the overall picture presented by the data.

If you have any concerns about your rights or how you are being treated please contact Eric Allen in the Office of Research and Compliance at UNCG at (336) 256-1482. Questions about this project or your benefits or risks associated with being in this study can be answered by Preston Yarborough who may be contacted at (336) 803-2149 or by e-mail at jpyarbor@uncg.edu.

Are there any benefits to me for taking part in this research study?

Participants in this study are interested in helping improve student /entry-level professional leadership skills. These participants might derive satisfaction from contributing to a process that can advance our understanding of leadership development.

Leaders and observers will have an opportunity to participate in a feedback session conducted with each participating institution's department of housing and residence life. This session will review macro-level trends seen across all participating institutions as well as trends unique within their specific institution.

Leaders attending this feedback session who were rated by at least 6 observers will receive an individualized feedback report. This report will aggregate responses from observers so they can have a better perspective of how their leadership style is perceived. Observers' anonymity is protected because this report contains only quantitative data and the data is aggregated such that a specific rater's feedback cannot be isolated or identified. A second, leaders-only debrief session will help leaders understand this individualized data and provide insight for future growth and development.

Observers (Resident Advisors) can benefit from the opportunity to voice an anonymous perspective of their Hall Director's leadership style. They also can benefit if their department of housing and residence life has a better understanding of the leadership behaviors and related conditions contributing to effective and ineffective leadership.

Are there any benefits to society as a result of me taking part in this research?

Leadership research has empirically demonstrated the need and value for multi-rater feedback data. Multi-rater feedback data is more reliable and has significantly greater predictive validity than self-reported leadership data. Very few multi-rater leadership instruments are appropriate for use in college settings. This study will produce a new multi-rater instrument that can measure participants' effective (or ineffective) use of leadership skills. This information may be used to inform leadership development training and to evaluate results of training initiatives.

Will I get paid for being in the study? Will it cost me anything?

There are no costs to you or payments made for participating in this study. There will be a drawing for two i-Pod Shuffles. There will be one winner from the Observer (RA) pool of participants and one winner from the Leader (HD) pool of participants.

How will you keep my information confidential?

During the study several measures have been enacted to preserve your confidentiality. First, when you submit your form the data will be protected by Secure Socket Layer (SSL) protection. This is the same security measures used to protect internet credit card transactions. Second, your identity and participation will not be revealed. You will be asked to enter your e-mail (minus the “@” symbol and the institutional address). For example, the e-mail jpyarbor@uncg.edu would be entered as jpyarbor. When you submit your form this information will be electronically transformed into a unique ID code. This code prevents your feedback data from having an outwardly identifiable name associated with it. There is a remote possibility someone with the key for the code could identify you as a participant. To minimize this risk only the researcher will possess a key to this code and only the researcher can access the key. The decoding key will be kept in a password-protected file on the researcher’s personal computer.

Absolute confidentiality of data provided through the Internet cannot be guaranteed due to the limited protections of Internet access. Please be sure to close your browser when finished so no one will be able to see what you have been doing.

Consent forms and data from this study will be kept for 3 years. After this duration, documents and related materials will be destroyed through the use of a paper shredder and/or will be electronically deleted.

What if I want to leave the study?

You have the right to refuse to participate or to withdraw at any time, without penalty. If you do withdraw, it will not affect you in any way. If you choose to withdraw, you may request that any of your data which has been collected be destroyed unless it is in a de-identifiable state.

What about new information/changes in the study?

If significant new information relating to the study becomes available which may relate to your willingness to continue to participate, this information will be provided to you.

Voluntary Consent by Participant:

By clicking the link below and submitting your responses to the survey form, you are implying that you have read this document and you fully understand the contents of this document and are openly willing consent to take part in this study. All of your questions concerning this study have been answered. By submitting your responses to the survey form, you are agreeing that you are 18 years of age or older and are agreeing to participate in this study described to you by Preston Yarborough.

Before entering the survey, you are encouraged to print this document in order to have a copy for your personal records.

[Click Here to Enter Survey Form](#)

APPENDIX U. INFORMED CONSENT, RESIDENT ADVISOR

UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

Section 1.05

CONSENT TO ACT AS A HUMAN

PARTICIPANT: LONG FORM

Project Title: Development and Validation of the Leadership Versatility Index for Students

Project Director: James Preston Yarborough

What is the study about?

This research study is focused on the development a multi-rater (360-degree) leadership assessment instrument (Leadership Versatility Index for Students—LVI-S) for students and entry-level professionals working in Housing and Residence Life. This instrument explores leadership versatility, or a leader's ability to effectively use a wide range of leadership behaviors. Few multi-rater leadership instruments have been developed specifically for collegiate audiences. The LVI-S will help address that need.

Why are you asking me?

You have been selected for participation in this pilot phase because you are serving as a Resident Advisor and have direct knowledge of leadership style of your supervising Hall Director.

What will you ask me to do if I agree to be in the study?

You will be asked to complete a 15 -20 minute on-line survey on your supervisor's leadership characteristics. Once you press the "submit" button at the end of the survey, the data will be transferred into a password-protected database.

What are the dangers to me?

Participation in this study is strictly voluntary. If you choose to participate, the information you provide will be strictly confidential and will not impact your relationship with Preston Yarborough, the UNCG Office of Leadership and Service-Learning, or your standing with your current department of employment. Should you choose not to participate, this decision will in no way affect your relationship or standing with Preston Yarborough or his employer, the Office of Leadership and Service-Learning at UNC-Greensboro. It will not affect your relationship with your current department of employment.

Participants may feel some risk rating the leadership characteristics and effectiveness of their supervisor. Several measures have been implemented to ease this concern. Hall Directors will not be informed which Resident Advisors participated in the study. Furthermore, individual observer responses will be kept confidential. Hall Directors that have 6 or more observers are eligible to receive a feedback report that summarizes the feedback from their observers, but any reported data will be presented in aggregate form (in other words, no individual-level responses will be reported). This makes it extremely difficult, if not impossible, to determine any single rater's scores.

Your participation can help us better understand the nature of effective leadership in residence halls. You will be helping to validate a new assessment tool that can enhance the leadership development and training of residence hall supervisors and RAs.

If you have any concerns about your rights or how you are being treated please contact Eric Allen in the Office of Research and Compliance at UNCG at (336) 256-1482. Questions about this project or your benefits or risks associated with being in this study can be answered by Preston Yarborough who may be contacted at (336) 803-2149 or by e-mail at jpyarbor@uncg.edu.

Are there any benefits to me for taking part in this research study?

Participants in this study are interested in helping improve student /entry-level professional leadership skills. These participants might derive satisfaction from contributing to a process that can advance our understanding of leadership development.

Leaders and observers will have an opportunity to participate in a feedback session conducted with each participating institution's department of housing and residence life. This session will review macro-level trends seen across all participating institutions as well as trends unique within their specific institution.

Leaders attending this feedback session who were rated by at least 6 observers will receive an individualized feedback report. This report will aggregate responses from observers so they can have a better perspective of how their leadership style is perceived. Observers' anonymity is protected because this report contains only quantitative data and the data is aggregated such that a specific rater's feedback cannot be isolated or identified. A second, leaders-only debrief session will help leaders understand this individualized data and provide insight for future growth and development.

Observers (Resident Advisors) can benefit from the opportunity to voice an anonymous perspective of their Hall Director's leadership style. They also can benefit if their department of housing and residence life has a better understanding of the leadership behaviors and related conditions contributing to effective and ineffective leadership.

Are there any benefits to society as a result of me taking part in this research?

Due to the pending retirement of the baby-boom generation workers, there will be a deficit of leaders available to fill an abundance of leadership positions. Consequently, we need to develop leaders earlier in their careers and we need to do so through effective methods. Collegiate leaders, and more specifically, paraprofessionals and professionals in residence life departments, are an important source of current and future leaders. But very few multi-rater leadership instruments are appropriate for use in college settings. This study will produce a new multi-rater instrument that can measure participants' effective (or ineffective) use of leadership skills. This information may be used to inform leadership development training and to evaluate results of training initiatives so that we can further enhance leadership development programs.

Will I get paid for being in the study? Will it cost me anything?

There are no costs to you or payments made for participating in this study. There will be a drawing for two i-Pod Shuffles. There will be one winner from the Observer (RA) pool of participants and one winner from the Leader (HD) pool of participants.

How will you keep my information confidential?

During the study several measures have been enacted to preserve your confidentiality. First, when you submit your form the data will be protected by Secure Socket Layer (SSL) protection. This is

the same security measure used to protect internet credit card transactions. Second, your identity and participation will not be revealed. You will be asked to enter your e-mail (minus the “@” symbol and the institutional address). For example, the e-mail jpyarbor@uncg.edu would be entered as jpyarbor. When you submit your form this information will be electronically transformed into a unique ID code. This code prevents your feedback data from having an outwardly identifiable name associated with it. There is a remote possibility someone with the key for the code could identify you as a participant. To minimize this risk only the researcher will possess a key to this code and only the researcher can access the key. The decoding key will be kept in a password-protected file on the researcher’s personal computer.

Absolute confidentiality of data provided through the Internet cannot be guaranteed due to the limited protections of Internet access. Please be sure to close your browser when finished so no one will be able to see what you have been doing.

Consent forms and data from this study will be kept for 3 years. After this duration, documents and related materials will be destroyed through the use of a paper shredder and/or will be electronically deleted.

What if I want to leave the study?

You have the right to refuse to participate or to withdraw at any time, without penalty. If you do withdraw, it will not affect you in any way. If you choose to withdraw, you may request that any of your data which has been collected be destroyed unless it is in a de-identifiable state.

What about new information/changes in the study?

If significant new information relating to the study becomes available which may relate to your willingness to continue to participate, this information will be provided to you.

Voluntary Consent by Participant:

By clicking the link below and submitting your responses to the survey form, you are implying that you have read this document and you fully understand the contents of this document and are openly willing consent to take part in this study. All of your questions concerning this study have been answered. By submitting your responses to the survey form, you are agreeing that you are 18 years of age or older and are agreeing to participate in this study described to you by Preston Yarborough.

Before entering the survey, you are encouraged to print this document in order to have a copy for your personal records.

[Click Here to Enter Survey Form](#)

APPENDIX V. INFORMED CONSENT SUPERVISOR OF HALL DIRECTOR

UNIVERSITY OF NORTH CAROLINA AT GREENSBORO

Section 1.06

CONSENT TO ACT AS A HUMAN

PARTICIPANT: LONG FORM

Project Title: Development and Validation of the Leadership Versatility Index for Students

Project Director: James Preston Yarborough

Participant's Name: _____

What is the study about?

This research study is focused on the development a multi-rater (360-degree) leadership assessment instrument (Leadership Versatility Index for Students—LVI-S) for students and entry-level professionals working in Housing and Residence Life. This instrument explores leadership versatility, or a leader's ability to effectively use a wide range of leadership behaviors. Few multi-rater leadership instruments have been developed specifically for collegiate audiences. The LVI-S will help address that need.

Why are you asking me?

You have been selected for participation in this pilot phase because you are serving as the supervisor of a Hall Director (or the equivalent position at your institution), have direct knowledge of the roles and responsibilities of this position, and direct knowledge of how Hall Directors have performed in this role.

What will you ask me to do if I agree to be in the study?

You will be asked to complete a very brief (3-item) multiple-choice questionnaire on the effectiveness of the Hall Directors you supervise. While you will need to complete a set of questions for each Hall Director you supervise, each set of questions should take less than 2 minutes to complete.

These ratings help to correct for a common form of bias in leadership research called same source bias. When behavioral ratings are collected from a source and then effectiveness ratings are taken from that same source, then the validity is limited by same source bias (Podsakoff & Organ, 1986). Collecting effectiveness ratings from the Hall Director's supervisor while collecting behavioral ratings from RAs reporting to the Hall Director helps control for same-source bias.

What are the dangers to me?

If you choose to participate, the information you provide will be strictly confidential and will not impact your relationship with Preston Yarborough, the Office of Leadership and Service-Learning, or your standing with your current department of employment. Should you choose not to participate, this decision will in no way affect your relationship or standing with Preston Yarborough or his employer, the Office of Leadership and Service-Learning at UNC-Greensboro. It will not affect your relationship with your current department of employment.

Participants may feel some risk at rating a person's effectiveness for the purposes of a research study. All participant data will be kept confidential and any reported data will be presented in aggregate form (in other words, no individual-level data will be reported). Your participation can

help us better understand the nature of effective leadership in residence halls. You will be helping to validate a new assessment tool that can enhance the leadership development and training of residence hall supervisors and RAs.

If you have any concerns about your rights or how you are being treated please contact Eric Allen in the Office of Research and Compliance at UNCG at (336) 256-1482. Questions about this project or your benefits or risks associated with being in this study can be answered by Preston Yarborough who may be contacted at (336) 803-2149 or by e-mail at jpyarbor@uncg.edu.

Are there any benefits to me for taking part in this research study?

Participants in this study are interested in helping improve student /entry-level professional leadership skills. These participants might derive satisfaction from contributing to a process that can advance our understanding of leadership development.

Participants in the study will have an opportunity to participate in a feedback session conducted with each participating institution's department of housing and residence life. This session will review macro-level trends seen across all participating institutions as well as trends unique within their specific institution. This session is intended to help participants and institutions transfer insights gained from this study into effective practice. This information may be used to inform leadership development training and to evaluate results of training initiatives so that you can further enhance leadership development within your institution.

Are there any benefits to society as a result of me taking part in this research?

Due to the pending retirement of the baby-boom generation workers, there will be a deficit of leaders available to fill an abundance of leadership positions. Consequently, we need to develop leaders earlier in their careers and we need to do so through effective methods. Collegiate leaders, and more specifically, paraprofessionals and professionals in residence life departments, are an important source of current and future leaders. But very few multi-rater leadership instruments are appropriate for use in college settings. This study will produce a new multi-rater instrument that can measure participants' effective (or ineffective) use of leadership skills. This information may be used to inform leadership development training and to evaluate results of training initiatives so that we can further enhance leadership development programs.

Will I get paid for being in the study? Will it cost me anything?

There are no costs to you or payments made for participating in this study.

How will you keep my information confidential?

During the study several measures have been enacted to preserve participant confidentiality. Hall Directors are aware that their supervisors are providing effectiveness ratings. Your ratings will not be shared with participants in this study and all data reported will be in aggregate form. Individual ratings will not be reported.

All data from the study will be recorded in a password-protected, secure database. This database does not contain any outwardly identifiable data because participant IDs are encoded.

Consent forms and data from this study will be kept for 3 years. After this duration, documents and related materials will be destroyed through the use of a paper shredder and/or will be electronically deleted.

What if I want to leave the study?

You have the right to refuse to participate or to withdraw at any time, without penalty. If you do withdraw, it will not affect you in any way. If you choose to withdraw, you may request that any of your data which has been collected be destroyed unless it is in a de-identifiable state.

What about new information/changes in the study?

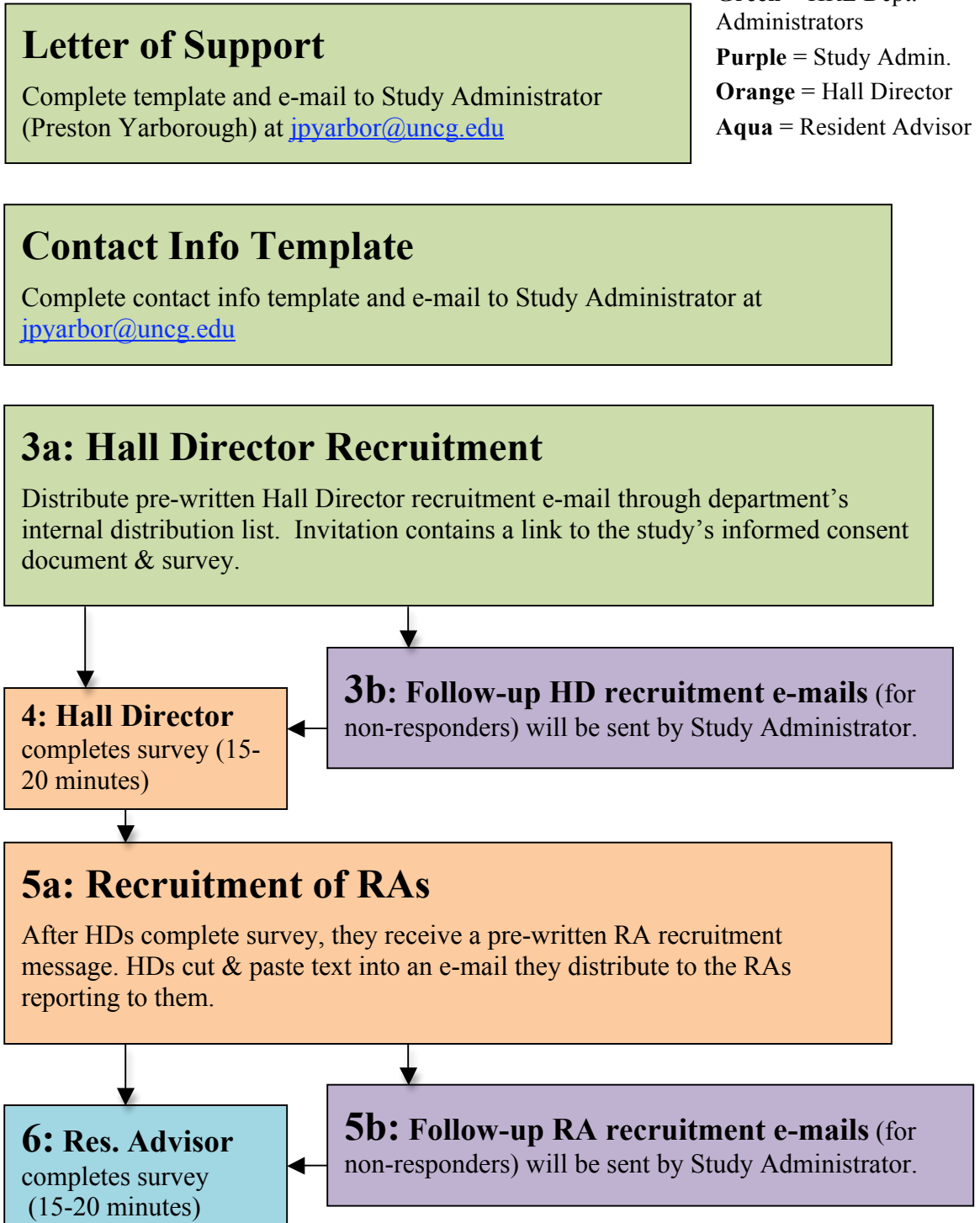
If significant new information relating to the study becomes available which may relate to your willingness to continue to participate, this information will be provided to you.

Voluntary Consent by Participant:

By signing this consent form you are agreeing that you have read and you fully understand the contents of this document and are openly willing consent to take part in this study. All of your questions concerning this study have been answered. By signing this form, you are agreeing that you are 18 years of age or older and are agreeing to participate, in this study described to you by Preston Yarborough.

Signature: _____ Date: _____

APPENDIX W. ADMINISTRATIVE PROCESSES AND PROCEDURES



APPENDIX X. EXPLANATION OF JOINT-STANDING SCORES

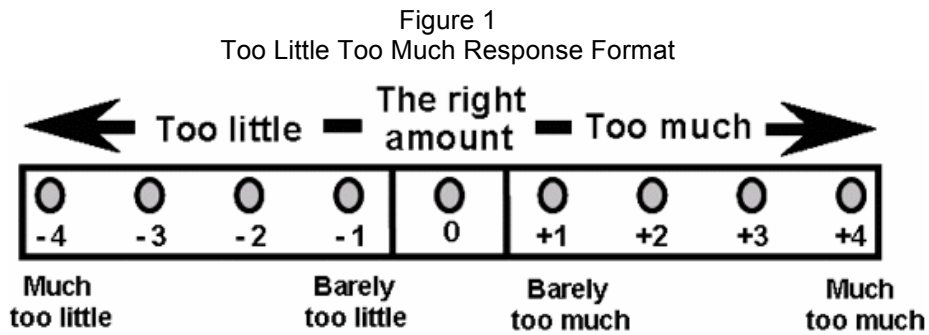
This appendix explains the conceptual purpose for joint-standing scores and how they are calculated. It is written with the presumption that a reader has familiarity with the conceptual structure of the LVI-S and the rationale behind the TLTM response scale. If needed, an explanation of versatility and the conceptual structure of the LVI-S has been elaborated in Chapters II and III.

The joint-standing variable is essential for measuring versatility. It calculates the leader's relative standing on complementary items, sub-dimensions, and dimensions. This "relative standing" refers to the leader's ability to dynamically balance complementary, but opposing behaviors. For example, consider the relationship between listening and speaking. Both are essential for communication—thus they complement one another. But both oppose each other as well—one cannot simultaneously listen and speak. The effective communicator seeks to establish a dynamic balance between listening the right amount and speaking the right amount. Thus, a joint-standing score for speaking and listening would measure observers' perceptions of the focal leader's ability to negotiate between listening and speaking.

The following explains how joint-standing scores were calculated at the item level, the sub-dimension level and the dimension level.

Item-level calculation. The most discrete level of measurement occurs between complementary item pairs. Responses to each item in a pair are recorded through the Too Little Too Much (TLTM) response format (Figure 1). Values on the TLTM format range

from -4 to +4, and the ideal score is the central value, 0, designated as “*the right amount*.”



The joint relationship of item pairs may be graphically illustrated by turning response scales for complementary items 90 degrees to one another. This forms a graph where the X and Y-axes range from -4 to +4, and intersect at the origin (0, 0) (see Figure 2).

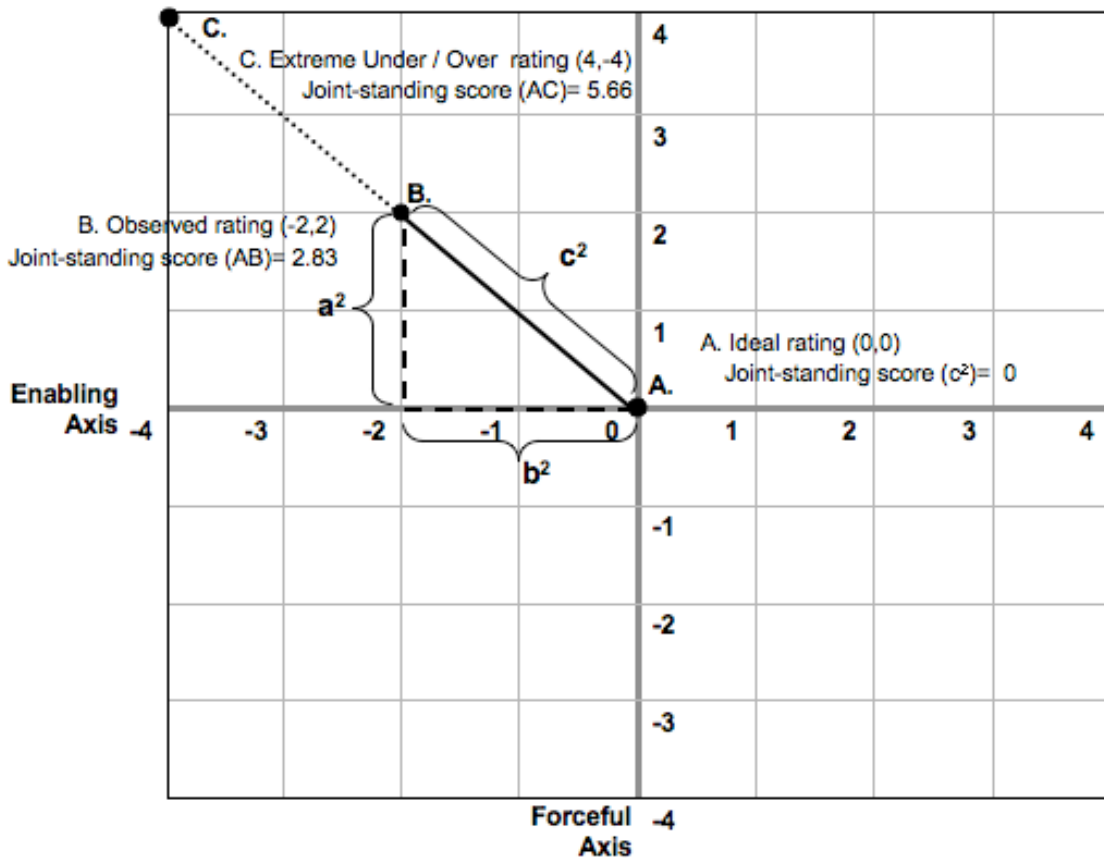
Observers’ responses may be plotted as a coordinate pair (an Enabling score plots on the X axis and a Forceful score plots on the Y axis).

The joint-standing score is calculated through the Pythagorean theorem ($a^2 + b^2 = c^2$). The ideal score on a matched pair of Forceful and Enabling items would be plotted at (0, 0), indicating both items were scored at “*the right amount*” (see Figure 2, point A).

Inserting these ratings into the Pythagorean formula produces a joint-standing score of zero, because $0^2 + 0^2 = 0^2$. This joint-standing score represents perfect versatility between the two items. Conversely, a score of (-4, 4), or any permutation of extreme positive or negative ratings, produces the least favorable joint-standing score (see Figure 1, point C). When inserted into the Pythagorean formula, these ratings yield a joint-standing score of

5.66, the square root of $(-4^2 + 4^2 = c^2)$. This joint-standing score (the distance of AC in Figure 1) represents the maximum distance from the ideal. Such a score represents extreme lopsidedness between the two items, or a complete absence of versatility with respect to an item pair. Joint-standing scores range between these two extremes. The Observed score $(-2, 2)$, designated by point B in Figure 1, represents an intermediate value between the ideal rating and the extreme rating. The distance of line AB is equal to 2.83; this result is 50 % less versatile than a leader with an ideal score on the same pair of items.

Figure 2
Calculation of Joint-standing Scores



Sub-dimension level calculation. Each sub-dimension in the LVI-S is composed of four items and each of these items is paired with a complement in the opposing sub-dimension. Thus, each sub-dimension has four joint-standing scores produced at the item level. These four joint-standing scores are averaged, producing a sub-dimension level joint-standing score. For example, the Takes Charge sub-dimension (Forceful) has four items paired with complements in the Empowers Others (Enabling) sub-dimension. The Takes Charge / Empowers Others joint-standing score is an average of the four item-level joint-standing scores (Ftc_1 / Ee_1; Ftc_2 / Ee_2; Ftc_3 / Ee3; Ftc_4 / Ee_4).

Duality level calculation. The duality level joint-standing score is an extension of the sub-dimension level calculations. Sub-dimension joint-standing scores are an average of the scores computed from the four item pairings within those sub-dimensions. The duality level joint-standing score is an average of the three sub-dimension-level joint-standing scores.

APPENDIX Y. LVI-S SUB-DIMENSION INTER-ITEM CORRELATIONS

Inter-item Correlations: Forceful Sub-dimensions

| | Ftc_1 | Ftc_2 | Ftc_3 | Ftc_4 | Fd_5 | Fd_6 | Fd_7 | Fd_8 | Fp_9 | Fp_10 | Fp_11 | Fp_12 |
|-------|-------|-------|-------|-------|------|------|------|------|------|-------|-------|-------|
| Ftc_1 | 1 | 0.53 | 0.15 | 0.28 | - | - | - | - | - | - | - | - |
| Ftc_2 | 0.53 | 1 | 0.31 | 0.34 | - | - | - | - | - | - | - | - |
| Ftc_3 | 0.15 | 0.31 | 1 | 0.27 | - | - | - | - | - | - | - | - |
| Ftc_4 | 0.28 | 0.34 | 0.27 | 1 | - | - | - | - | - | - | - | - |
| Fd_5 | - | - | - | - | 1 | 0.41 | 0.18 | 0.26 | - | - | - | - |
| Fd_6 | - | - | - | - | 0.41 | 1 | 0.50 | 0.43 | - | - | - | - |
| Fd_7 | - | - | - | - | 0.18 | 0.50 | 1 | 0.31 | - | - | - | - |
| Fd_8 | - | - | - | - | 0.26 | 0.43 | 0.31 | 1 | - | - | - | - |
| Fp_9 | - | - | - | - | - | - | - | - | 1 | 0.58 | 0.40 | 0.33 |
| Fp_10 | - | - | - | - | - | - | - | - | 0.58 | 1 | 0.38 | 0.34 |
| Fp_11 | - | - | - | - | - | - | - | - | 0.40 | 0.38 | 1 | 0.64 |
| Fp_12 | - | - | - | - | - | - | - | - | 0.33 | 0.34 | 0.64 | 1 |

Ftc_ = Takes Charge

Fd_ = Declares

Fp_ = Pushes

Inter-item Correlations: Enabling Sub-dimensions

| | Ee_13 | Ee_14 | Ee_15 | Ee_16 | El_17 | El_18 | El_19 | El_20 | Es_21 | Es_22 | Es_23 | Es_24 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Ee_13 | 1 | 0.49 | 0.21 | 0.32 | - | - | - | - | - | - | - | - |
| Ee_14 | 0.49 | 1 | 0.37 | 0.44 | - | - | - | - | - | - | - | - |
| Ee_15 | 0.21 | 0.37 | 1 | 0.45 | - | - | - | - | - | - | - | - |
| Ee_16 | 0.32 | 0.44 | 0.45 | 1 | - | - | - | - | - | - | - | - |
| El_17 | - | - | - | - | 1 | 0.57 | 0.47 | 0.50 | - | - | - | - |
| El_18 | - | - | - | - | 0.57 | 1 | 0.53 | 0.36 | - | - | - | - |
| El_19 | - | - | - | - | 0.47 | 0.53 | 1 | 0.51 | - | - | - | - |
| El_20 | - | - | - | - | 0.50 | 0.36 | 0.51 | 1 | - | - | - | - |
| Es_21 | - | - | - | - | - | - | - | - | 1 | 0.49 | 0.31 | 0.24 |
| Es_22 | - | - | - | - | - | - | - | - | 0.49 | 1 | 0.46 | 0.40 |
| Es_23 | - | - | - | - | - | - | - | - | 0.31 | 0.46 | 1 | 0.34 |
| Es_24 | - | - | - | - | - | - | - | - | 0.24 | 0.40 | 0.34 | 1 |

Ee_ = Empowers Others

El_ = Listens

Es_ = Supports

APPENDIX Z. LVI-S CONCEPTUAL STRUCTURE DIAGRAM

